Winter 12-11-2012

Biophysical Characterization of Synthetic Imidazole and Pyrrole Containing Analogues of Netropsin and Distamycin that Target Specific DNA Sequences for the Treatment of Various Diseases

Joseph P. Ramos
Georgia State University

BIOPHYSICAL CHARACTERIZATION OF SYNTHETIC IMIDAZOLE AND PYRROLE CONTAINING ANALOGUES OF NETROPSIN AND DISTAMYCIN THAT TARGET SPECIFIC DNA SEQUENCES FOR THE TREATMENT OF VARIOUS DISEASES

by

JOSEPH PAUL RAMOS

Under the Direction of Dr. W. David Wilson

ABSTRACT

The development of small-molecules which target nucleic acids, more specifically the minor groove of DNA, in a sequence specific manner and control gene expression are currently being investigated as potential therapeutic compounds for the treatment of various diseases, including cancer, as well as viral and bacterial infections. The naturally occurring compounds netropsin and distamycin have been shown to demonstrate antitumor and antibacterial properties. Currently, there are synthetic efforts to create pyrrole and imidazole-containing polyamide derivatives of netropsin and distamycin that show potential as medicinal agents. Synthetic pyrrole and imidazole-containing polyamides are potentially useful for targeting and modulating the expression of genes, including those associated with cancer cell growth.

The key challenges that must be overcome to realize this goal of using synthetic polyamides in the treatment of disease are the development of polyamides with low molar mass so the molecules can readily diffuse into cells and concentrate in the nucleus. In addition, the molecules must have appreciable water solubility, bind DNA sequence specifically, and with

high affinity. As part of a systematic study within the authors' laboratory, our goal is to develop polyamides which can be synthesized readily yet possess excellent sequence specificity, stronger binding affinity, high solubility in biological media and enhanced cell penetration and nuclear localization properties.

There is a need to develop a library of modified polyamides which target DNA and exhibit improved biological properties. The present study is a systematic examination of the binding properties of various modified synthetic polyamide compounds. The synthetic polyamide derivatives presented have more potential as therapeutic candidates over other synthetic polyamides because of their increased water solubility, smaller molecular weights, and molecular design, thus, allowing them to penetrate into cells and localize in the nucleus.

INDEX WORDS*: DNA, minor groove, polyamides, netropsin, distamycin, imidazole, pyrrole, sequence specificity, gene control, surface plasmon resonance, isothermal titration calorimetry*

BIOPHYSICAL CHARACTERIZATION OF SYNTHETIC IMIDAZOLE AND PYRROLE CONTAINING ANALOGUES OF NETROPSIN AND DISTAMYCIN THAT TARGET SPECIFIC DNA SEQUENCES FOR THE TREATMENT OF VARIOUS DISEASES

by

JOSEPH PAUL RAMOS

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2012

BIOPHYSICAL CHARACTERIZATION OF SYNTHETIC IMIDAZOLE AND PYRROLE CONTAINING ANALOGUES OF NETROPSIN AND DISTAMYCIN THAT TARGET SPECIFIC DNA SEQUENCES FOR THE TREATMENT OF VARIOUS DISEASES

by

JOSEPH PAUL RAMOS

Committee Chair: Dr. W. David Wilson

Committee: Dr. Markus W. Germann

Dr. Dabney W. Dixon

Electronic Version Approved:

Office of Graduate Studies

Colleges of Arts and Sciences

Georgia State University

August 2012

DEDICATIONS

…This work is dedicated to my family, most importantly, my parents Karla and Barney, whose unwavering love and support have made this all possible! You two are the greatest parents a person could have! I love you both with every fiber of my being!

…Also, I would like to dedicate this work to those people working to help cure disease and those affected by disease! Whether your contribution is large or small, it is because of people like you, people suffering from illness have hope! May we all work towards improving the quality of life of those affected and, if even just for a day, extend the quality of their lives! Life is beautiful and precious; let us NEVER forget that, nor what we are working for!

ACKNOWLEDGEMENTS

It would be truly impossible to list all the wonderful people who have helped and/or influenced me in various ways throughout the complex process of obtaining my doctoral degree, never the less, I must acknowledge a certain few.

First and foremost, I feel it is imperative, an absolute must, and of the utmost importance to recognize one of the most humble, kindhearted, genuine, intelligent, and enthusiastic persons I have ever had the privilege of knowing, Dr. W. David Wilson. Without Dr. Wilson's unwavering support, guidance, understanding, and belief in my abilities, I would not have had the opportunity to finish the very daunting task of completing my doctoral degree. Dr. Wilson's ever-lasting love and enthusiasm for science is without a doubt, absolutely unrivaled! Not only has he been an outstanding advisor and mentor, but he has also become a very well-respected and life-long friend, someone whom I admire greatly! Along with all the science he has taught me, even more importantly he has taught me to be a man, a good friend and colleague, and a good human-being. I have the utmost gratitude and am forever in debt to Dr. Wilson for everything he has done for me!

Secondly, I must give a very special thanks to my committee members, Drs. Markus W. Germann and Dabney W. Dixon. I am forever indebted to both of them for their constructive criticisms of my work and research, for their excellent direction and input, and most importantly for accepting to be my committee members and finding my research and work worthy of a PhD. Through being both members of my dissertation committee and professors of whom I have taken numerous courses, they have challenged me, pushed me to think critically and outside the box, and shared their wealth of knowledge with me! For this, I am grateful!

I would also like to thank all the members of the Wilson Laboratory, both past and present. Without their support and them laying the foundation for my work, none of this would have been possible. I would like to specifically thank Drs. Yang Liu, Rupesh Nanjunda, and Manoj Munde for sharing their knowledge of SPR with me and for taking the time, as well as having the patience, to teach me this brilliant technique! Drs. Liu and Munde were always more than willing to take the time to answer my questions, show me how to operate and take care of the SPR instruments, and engage in discussions regarding our work together, and they did this without hesitation! For that I am eternally grateful.

Along with members of the Wilson laboratory, this research could not have been completed without the collaboration of Dr. Moses Lee and Dr. James Baskin. Dr. Moses Lee was gracious enough to provide the majority of the compounds studied in this research. His youthful attitude and love of chemistry is only matched by that of Dr. Wilson. Dr. Lee and his group's synthetic efforts and contributions to the field of polyamides are without comparison! Without them, this work would not have been possible!

Without funding, the work presented in this dissertation would not have been possible. Therefore, I would like to thank NIH and NSF for funding the projects presented here. Also, I would like to thank the Georgia State University Department of Chemistry for providing four years of tuition waivers, as well as my first year stipend, and the Molecular Basis for Disease for funding my stipend the last three years of my doctoral research.

Lastly, and of the greatest significance, I would like to thank my family, my brother, Anthony, and his wife, Danyell, who have always believed in and supported me, and most importantly, my parents, Karla and Barney! Their love has been untiring and steadfast in

supporting my quest for knowledge and the pursuit of not only my doctoral degree but every goal I have set out to accomplish in life. It is because of a work ethic and determination instilled in me by them, that I have been able to accomplish all that I have. Their belief in me and what I want to accomplish in life is something that inspires me every day! Along with all the love and emotional support, they have sacrificed and supported me financially throughout this endeavor, and without that, I would not have been able to achieve such a goal. My parents have made my education a priority, so that I may live a life I can be proud of. They have taught me the importance of honesty, humility, hard work and love! Words simply cannot express how much their love and support has meant to me! I love you all more than words could ever possibly describe! You guys are my world, my everything! Thank you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

LIST OF ABBREVIATIONS

Im	*N*-methylimidazole
Py	*N*-methylpyrrole
f	formamido
Hp	3'-hydroxy-1 H-pyrrole
DNA	deoxyribonucleic acid
A	adenine
T	thymine
G	guanine
C	cytosine
MG	minor groove
Oligo	oligonucleotide
nt	nucleotide
ss	single-stranded
ds	double-stranded
HP DNA	hairpin deoxyribonucleic acid
ITC	isothermal titration calorimetry
SPR	surface plasmon resonance
RU	response units
CD	circular dichroism
UV	ultraviolet
Vis	visible
T_m	thermal melting temperature
mp	melting point

ppm	parts per million
CCL	cacodylic acid
HEPES	N-[2-Hydroxyethyl]piperazine-N´-[2-enthanesulfonic acid]
EDTA	ethylenediaminetetraacetic acid
NIH	National Institutes of Health
NSF	National Science Foundation

CHAPTER 1:

AN INTRODUCTION TO AND HISTORY OF POLYAMIDES: NETROPSIN AND DISTAMYCIN, NATURALLY OCCURING POLYAMIDES THAT TARGET THE MINOR GROOVE OF DNA IN A SEQUENCE SPECIFIC MANNER AND DEMONSTRATE BIOLOGICAL ACTIVITY

1.1 Discovery of netropsin and distamycin

Netropsin is a naturally occurring oligopeptide first discovered and isolated by Finlay *et al.* in 1951 (*1*). This oligopeptide has been of significant interest over the last sixty years because of its biological activity. Netropsin has shown both antibacterial and antiviral activity and is classified as a pyrrole (Py)-containing amide antibiotic, although it is not used clinically because of toxicity. It was isolated from the actinobacterium *Streptomyces netropsis* and given the name "netropsin" (*1*). Netropsin small molecules consist of two pyrrole rings and two terminal amidine groups linked together by amide (peptide) bonds, as shown in Figure 1A.

Distamycin is also a naturally occurring oligopeptide discovered in 1958 by Arcamone *et al.* and isolated from an actinomycete bacterium, *Streptomyces distallcius*, of the same *Streptomyces* genus as netropsin. Distamycin was discovered as a byproduct from the fermentation of *S. distallcius*. Distamycin, like netropsin, has demonstrated antiviral and antibiotic properties, as well as antitumor characteristics. Distamycin is very similar in structure to netropsin although it does have an additional pyrrole ring, thus, being comprised of three pyrrole (Py) rings, each connected by a peptide bond, one terminal amidine (instead of two like netropsin), and a terminal formamido (f) group, as shown in Figure 1B. Distamycin is classified as a pyrrole-containing amidine antibiotic but is also not used clinically due to its cytotoxicity.

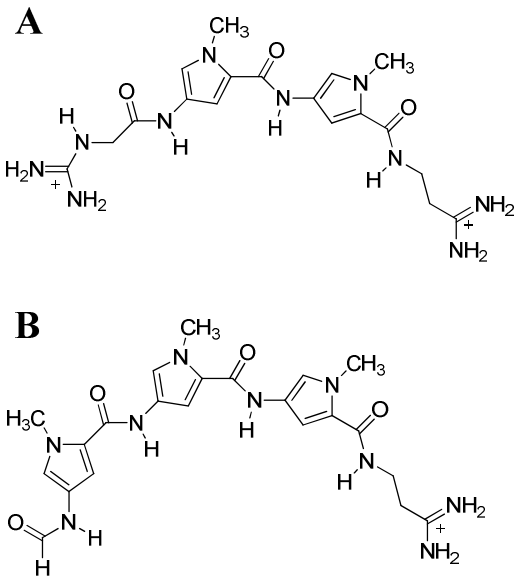


Figure 1. Chemical structure of the polyamide netropsin (A) and distamycin (B).

1.2 Polyamides as DNA minor groove binders

The major groove of DNA is most often the binding site for protein-DNA interactions. This is due to the diversity and accessibility of the functional groups of the base pairs to protein structures. However, this is not usually the case with small molecules, such as netropsin and distamycin which bind to the minor groove. The two grooves differ in electrostatic potential, hydration, hydrogen bonding, and steric properties (*2*). Moreover, they also have very different microstructures and conformations which are highly dependent upon base pair composition. For example, long consecutive runs of adenosine bases are known to exhibit a very narrow minor groove when compared to the minor groove of a stretch DNA that consists of guanine and cytosine base pairs (*2*). It is through these differences that the minor groove has become an attractive target for small molecules. Polyamide small molecules, such as netropsin and

distamycin, preferentially bind to the minor groove of DNA in a sequence specific manner. It has been shown that netropsin and distamycin preferentially bind to AT-rich sequences. The narrower minor groove enhances electrostatic interactions between the small molecule and the minor groove.

Netropsin and distamycin are comprised of multiple simple aromatic ring systems, linked together by bonds which have torsional freedom. The torsional freedom of these bonds allows them to adopt a crescent shape which is a typical, yet a key structural property of minor groove binding (*2*). The torsional freedom and crescent shape adopted by the small molecule allows it to twist in such a way that it becomes isohelical with the DNA minor groove. Because the polyamide and DNA are isohelical, contacts between the two are optimized (*2*).

The narrowness of AT-rich minor groove sequences allow for the optimal van der Waals interactions between the crescently shaped polyamide and the helical chains that define the walls of the minor groove. Additionally, hydrogen bonding between bound ligands and the base pairs found at the floor of the minor groove creates both sequence specificity and stability of the complex. A-T base pairs are able to accept hydrogen bonds from the small molecule through the C2-carbonyl oxygen of thymine and the N3 nitrogen of adenine. Although there are similar functional groups found on C-G bases, the exocyclic amine at the N3 position of guanine and the exocyclic oxygen at the O2 position of cytosine form a hydrogen bond and create a steric hindrance, and thus, the small molecules netropsin and distamycin bind specifically to AT-rich sequences. The hydrogen bond between the amine group of guanine and the carbonyl oxygen of cytosine extends into the minor groove also, adding to the steric hindrance and not allowing penetration in to the minor groove by netropsin or distamycin (*3*). It has also been shown that the electrostatic potential of the minor groove of AT-rich DNA sequences is much greater than

that of the minor groove of GC-rich DNA sequences, therefore, the favorable binding to AT-rich sequences by positively charged small molecules, such as netropsin and distamycin, is favorable. The ability of the exocyclic amine to become a hydrogen bond donor can be exploited for sequence recognition and enhanced binding affinity by modified polyamides. This concept will be addressed in section 1.4.

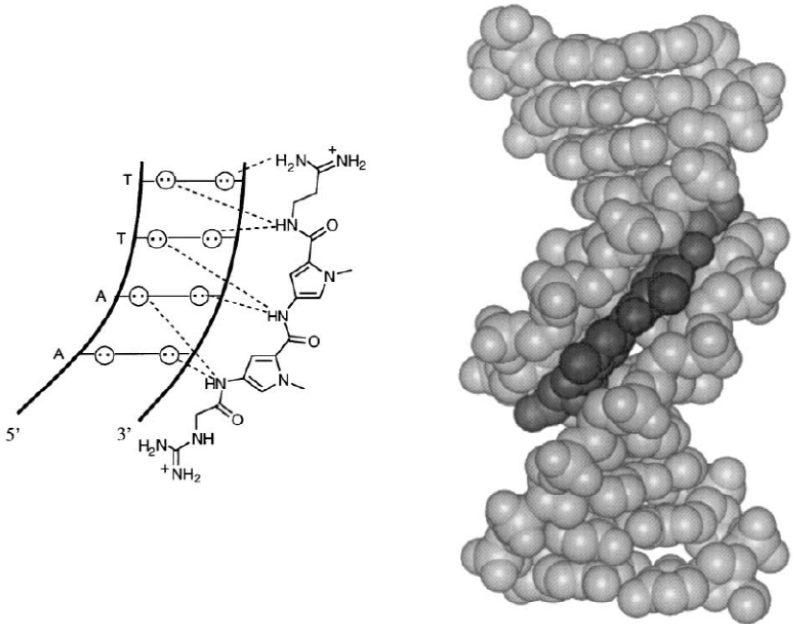


Figure 2. X-ray crystallographic structure of 1:1 netropsin-DNA complex (*4*).

The first crystal structure of a netropsin-DNA complex with 5'-(dCGCGAATTCGCG)2-3' was solved by Dickerson *et al.* in 1985 (*5*). The complex showed a single netropsin molecule bound to the AATT-site, as was expected. The three amide –NH groups point towards the floor of the minor groove forming bifurcated hydrogen bonds with the O2 of thymine and N3 of adenine, as shown in Figure 2. van der Waals contacts are formed between the tightly bound polyamide and the atoms which comprise the sugar-phosphate walls of the groove, creating a parallel plane between the pyrrole rings of netropsin and the walls of the groove. The interaction of the two cationic amidine groups and the N3 of the 5' adenine is favorable and a key component of complex formation. The methyl groups at the N5 position of the pyrrole rings point into the solution, resulting in an ability for the compound to penetrate into the minor groove.

The crystal structure for distamycin has several key features very similar to that of netropsin. However, an NMR study by Wemmer *et al.* in 1990 showed distamycin bound as a stacked homodimer to the sequence 5'-CGCGAAATTTCGCG-3'(Figure 3) (*3, 6-8*). The additional two base pairs (compared to the AATT site) in the central recognition site allowed for the "longer" distamycin molecule and could accommodate the stacked dimer even in the narrower AT-rich minor groove. Data from this study showed that the 2:1 distamycin-DNA complex was favored for binding sites six base pairs or longer, and the 1:1 complex for a binding site of four base pairs (5'-AATT-3').

Two distamycin molecules bind to the minor groove of DNA in the same fashion as netropsin. The amide hydrogens of the distamycin molecules face towards the floor of the minor groove, forming bifurcated hydrogen bonds with the O2 of thymine and N3 of adenine. The two ligands stack in a "staggered" manner in which the two positive charges of the amidine groups

are at opposite ends (antiparallel), thus maximizing the distance between the charges and reducing the electrostatic repulsion. Because netropsin has two positive charges on the terminal amidine groups, one on each end, the electrostatic repulsion would be too great for netropsin to bind as a homodimer and this explains why the observed stoichiometry of netropsin-DNA complexation is 1:1 and not 2:1. In order to accommodate the two distamycin molecules, the minor groove must widen. The structural change as a result of 2:1 distamycin-DNA complexation causes a positive, unfavorable overall free energy contribution. Although the widening of the groove is unfavorable, it is compensated by a highly favorable hydrophobic interaction that is a direct result of stacking and binding two ligands within one binding site of the minor groove.

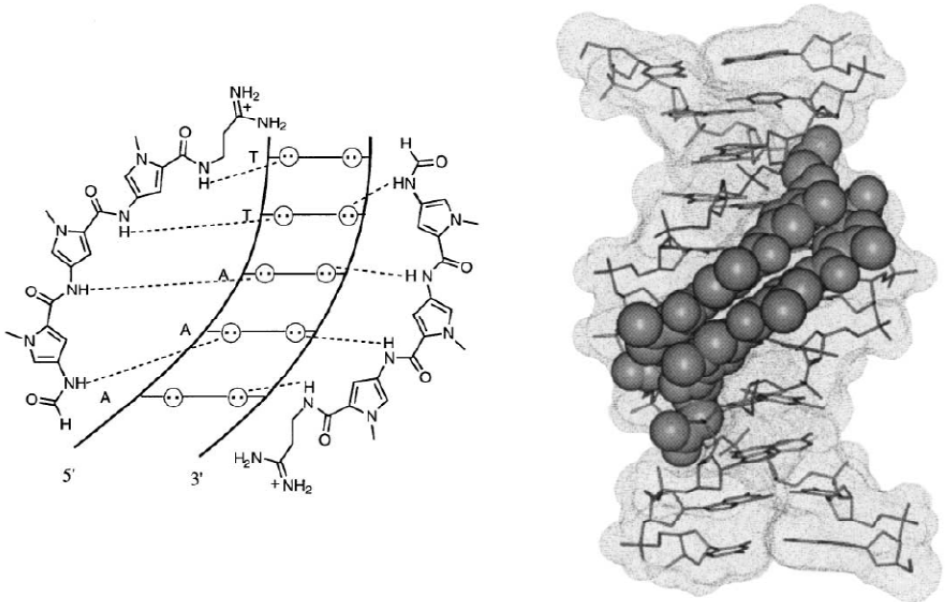


Figure 3. NMR structure of 2:1 distamycin-DNA complex (*4*).

1.3 Netropsin and distamycin interactions: A look at energetics

The thermodynamic and binding properties of both netropsin and distamycin to DNA have been studied extensively. Because of advancements in instrumentation over the last twenty years, enthalpic and entropic contributions to the overall free energy have been measured, giving a more complete thermodynamic profile and defining the driving force behind complexation. The 2:1 binding of distamycin to 5'-AAATT-3' compared to the 1:1 complex formation of netropsin and 5'-AAATT-3' have been examined and compared (*9*). Both netropsin and distamycin have favorable binding enthalpy contributions and are exothermic processes. Drug binding is most often associated with a favorable enthalpic binding process (*2, 3*). The enthalpy for netropsin binding to 5'-AAATT-3' is ~-31 kJ/mol. The enthalpy for binding two distamycin molecules to the same DNA sequence are ~-52 kJ/mol for the first molecule and ~79 kJ/mol for the second (*2*). K_{eq} for netropsin is ~10^8 M^{-1} (*10*) and K_1 and K_2 for distamycin are ~10^7 M^{-1} and 10^6 M^{-1}, respectively. Because K_2 is ~10 times less than K_1, the binding of two distamycin molecules to one binding site is a negatively cooperative process for five AT base pairs.

The binding process for netropsin and the first distamycin compound are very similar in nature. Both have similar binding affinities and are both enthalpically driven, although binding of the first distamycin molecule has a larger enthalpic contribution. Both polyamides have very similar hydrogen bonding patterns which could explain why the binding enthalpies are similar for both complexes. The presence of an additional cationic group for netropsin compared to distamycin allows for more favorable electrostatic interactions for netropsin. However, distamycin has an additional pyrrole ring which allows for more favorable hydrophobic interactions, and thus compensates for the decrease in electrostatic interactions. Although the enthalpies of binding for netropsin and distamycin are very similar, the entropic contributions to

the overall Gibbs free energy are opposite for the two polyamides. Netropsin has a favorable entropy whereas distamycin has an unfavorable entropy. The differences in entropy are attributed to changes in hydration; water is released or taken up upon binding, and the release of counter ions.

1.4 Synthetic derivatives of netropsin and distamycin

Since their discovery sixty years ago, netropsin and distamycin have served as templates for the design of synthetic polyamides. In the mid-late 1980's, Lown (*11*) and Dickerson *et al.* (*12*), developed synthetic polyamides, lexitropsins, which could target G-C and C-G sequences (*13*). This was accomplished by substituting an *N*-methyl pyrrole (Py) with an *N*-methyl imidazole (Im). The additional nitrogen in the imidazole ring acts as a hydrogen bond acceptor and can hydrogen bond with the exocyclic amino group of guanine. Targeting of G-C and C-G sequences through the use of Im rings developed by Lown and Dickerson were exploited by Dervan, who eventually developed a set of binding rules which are outlined in Figure 4 (*14-18*). Figure 4 shows a schematic view of binding the synthetic polyamide Im-Py-Py to the sequence 5'-TGTCA-3' (*4*). In theory, this concept allowed for the programmable targeting of any specific DNA sequence. The idea of programmably targeting specific DNA sequences led to an explosion in synthetic efforts of numerous laboratories throughout the scientific community.

The two differences between netropsin and distamycin are: 1) netropsin has two pyrrole rings and distamycin has three, and 2) netropsin has two positively charged termini groups and distamycin has one cationic amidine terminus and a formamido (f) group at the other terminus. Most synthetic efforts have left off the terminal formamido group because of synthetic difficulties (*19*). The influence of the terminal formamido on stacking and binding was formally

addressed and studied by Lacy *et al.* in 2001 (*20*). The addition of a terminal formamido led to a staggered, antiparallel stacking formation and a 10-100 fold increase in binding affinity depending on the synthetic polyamide. This systematic study revealed the importance of the terminal formamido group and its influence on the mode of stacking, which directly impacts the binding affinity and kinetics of complexation (*20*). The study also showed the presence of a terminal formamido group had no advantage in terms of sequence specificity.

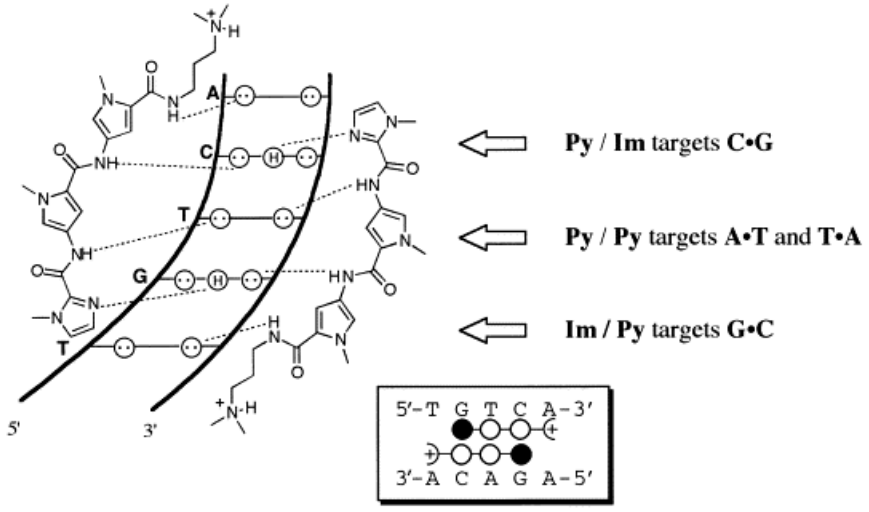


Figure 4. Schematic representation of G-C recognition of Im-Py and the binding of Im-Py-Py to 5'-TGTCA-3'. The dashed lines represent hydrogen bonds between the synthetic polyamide and its recognition sequence (*4*). The black circles of the inset signify Im rings, the open circles Py rings and the crescent shape + are the positively charged tail region.

From the development of Dervan's pairing rules, numerous different types of polyamides have been developed to optimize both binding affinity and sequence specificity. Upon the discovery of the 2:1 complex formation of distamycin with 5'-AAATTT-3' through Wemmer's NMR structure, synthetic developments have been made to covalently link the two dimers. This led to the development of hairpin polyamides, where two polyamides are linked from the 3' terminus (end) of one to the 5' terminus (tail) of the other, H-pin polyamides, where two polyamides are linked via N5 position of both central Py or Im rings, and cyclic polyamides, where two polyamides are linked from both the 5' and 3' ends to the opposing 3' and 5' ends of the second polyamide molecule. The lengths of the linking moieties have been well studied for all three polyamide derivatives, as have the addition of functional groups on the linkers (*21-24*).

1.5 Polyamides control gene expression: potential as therapeutics to treat various diseases

Since their discovery and isolation, netropsin and distamycin have been known to have biological activity as antibiotic, antiviral and antitumor therapeutics, leading to the hypothesis that these two molecules are synthesized naturally by streptomycete strains as a potential defense mechanism (*3*). Though too toxic for clinical use, these compounds have been used to elucidate the origins of sequence specificity and biological activity. These molecules are thought to interfere with topoisomerase in some fashion, although their true mechanism is not completely understood. Excitingly, synthetic Py- and Im-containing polyamides have not shown the same levels of toxicity as netropsin and distamycin.

The first publication to report the biological activity of a synthetic polyamide *in vivo* was by Gottesfeld in *Nature* in 1997 (*25*). The authors reported the activity of a hairpin polyamide, ImPyPyPy-γ-ImPyPyPy-*β*-Dp, where γ represents the γ-aminobutyric acid linker, *β* represents

the $(CH_2)_3$ alkyl linker which links the terminal secondary amine and the last pyrrole ring. This synthetic hairpin polyamide binds to the sequence 5'-TAGTACT-3', which is found within the promoter site of the TFIIA gene, and causes inhibition of the 5S RNA through direct competition. *In vivo* studies showed the presence of the hairpin polyamide led to a reduction in 5S RNA by interfering with the transcriptional process of the TFIIA gene. Although a debate remains about the viability of larger hairpin polyamides and their ability to permeate into the cells and localize in the nucleus, Gottesfeld *et al.* presented a strong first argument for the use of polyamides as therapeutics showing direct evidence addressing these issues (*3, 25*).

Modulation of gene expression, specifically the inverted 5'-CCAAT-3' of the topoisomerase IIα (topoIIα) gene, through the use of polyamide-DNA interactions has been of interest in the Lee laboratory (*26*). TopoIIα plays a role in many cell processes, including DNA replication and mitosis, and is a target of the anticancer drugs etoposide and doxorubicin (*27-29*). TopoIIα is down regulated in confluent cancer cells, and thus, these cells become resistant to treatment using the anticancer agents etoposide and doxorubicin. The nuclear factor-Y (NF-Y) protein is responsible for inhibition of gene expression at confluence. NF-Y binds to the inverted CCAAT-box 2 (ICB2, there are 5 total in the promoter region). Polyamide binding to ICB2 blocks the site and NF-Y cannot bind to it, leading to the upregulation of topoIIα expression and making them susceptible to treatment with etoposide and doxorubicin (*26, 30*). Lee's laboratory has shown that the synthetic polyamide f-PyImPy binds to the sequence 5'-TACGAT-3' which flanks the ICB2 binding site and inhibits the binding of NF-Y (*26*).

1.6 Objectives of this Dissertation

Although synthetic polyamides have shown promise as potential therapeutics, there is still a need to understand the fundamentals of their biological activity. Modification and modulation of gene expression by small molecules is still not completely understood. There has to be a fundamental understanding of how modifying synthetic polyamides affects polyamide-DNA complex formation, in terms of sequence specificity and binding affinity. This dissertation presents five publications which systematically characterize the effects of synthetically modified polyamides on binding affinity and sequence selectivity.

References

1. Finlay, A. C., Hochstein, F. A., Sobina, B. A., and Murphy, F. X. (1951) Netropsin, a new antibiotic produced by a Streptomyces, *J. Am. Chem. Soc. 73*, 341-343.
2. Blackburn, G. M., and Gait, M. J. (1997) *Nucleic Acids in Chemistry and Biology*, 2nd ed., Oxford University Press, New York.
3. Wemmer, D. E. (2001) Ligands recognizing the minor groove of DNA: Development and applications, *Biopolymers 52*, 197-211.
4. Dervan, P. B. (2001) Molecular recognition of DNA by small molecules, *Bioorg. Med. Chem. 9*, 2215-2235.
5. Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., and Dickerson, R. E. (1985) Binding of an antitumor drug to DNA, netropsin and C-G-C-G-A-A-T-T-BrC-G-C-G, *J. Mol. Biol. 183*, 553-563.
6. Pelton, J. G., and Wemmer, D. E. (1990) Binding modes of distamycin A with d(CGCAAATTTGCG)2 determined by two-dimensional NMR, *J. Am. Chem. Soc. 112*, 1393-1399.
7. Coll, M., Aymami, J., van der Marel, G. A., van Boom, J. H., Rich, A., and Wang, A. H. (1989) Molecular structure of the netropsin-d(CGCGATATCGCG) complex: DNA conformation in an alternating AT segment, *Biochemistry 28*, 310-320.
8. Coll, M., Fredrick, C. A., Wang, A. H.-J., Rich, A. P., and Natl Acad Sci USA 1987. (1987) A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin., *Natl Acad Sci USA 84*, 8385–8389.
9. Rentzeperis, D., Marky, L. A., Dwyer, T. J., Geierstanger, B. H., Pelton, J. G., and Wemmer, D. E. (1995) Interaction of minor groove ligands to an AAATT/AATTT site: Correlation of thermodynamic characterization and solution structure, *Biochemistry 34*, 2937-2945.
10. Nguyen, B., Tanious, F., and Wilson, W. D. (2006) Biosensor-surface plasmon resonance: Quantitative analysis of small molecule–nucleic acid interactions, *Methods 42*, 150-161.
11. Lown, J. W., Krowicki, K., Bhat, U. G., Skorobogaty, A., Ward, B., and Dabrowiak, J. C. (1986) Molecular recognition between oligopeptides and nucleic acids: novel imidazole-containing oligopeptides related to netropsin that exhibit altered DNA sequence specificity, *Biochemistry 25*, 7408–7416.
12. Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., and Dickerson, R. E. (1985) The molecular origin of DNA-drug specificity in netropsin and distamycin, *Proc. Natl. Acad. Sci. U S A 82*, 1376-1380.
13. Lee, M., Hartley, J. A., Pon, R. T., Krowicki, K., and Lown, J. W. (1988) Sequence spcific molecular recognition by a monocationic lexitropsin of the DNA d(CATGGCCATG): structural and dynamic aspects deduced from high field proton NMR studies, *Nucleic Acids Res. 16*, 665-684.
14. Mrksich, M., and Dervan, P. B. (1993) Antiparalllel side-by-side heterodimer for sequence-specific recognition in the minor groove of DNA by a distamycin/1-Methylimidazole-2-carboxamide-netropsin pair, *J. Am. Chem. Soc. 115*, 2572-2576.

15. Geierstanger, B. H., Jacobsen, J. P., Mrksich, M., Dervan, P. B., and Wemmer, D. E. (1994) Structural and dynamic characterization of the heterodimeric and homodimeric complexes of distamycin and 1-methylimidazole-2-carboxamide-netropsin bound to the minor groove of DNA, *Biochemistry 33*, 3055-3062.
16. Geierstanger, B. H., Mrksich, M., Dervan, P. B., and Wemmer, D. E. (1994) Design of a G.C-specific DNA minor groove-binding peptide, *Science 266*, 646-650.
17. Kielkopf, C. L., Baird, E. E., Dervan, P. B., and Rees, D. C. (1998) Structural basis for G.C recognition in the DNA minor groove, *Nat. Struct. Biol. 5*, 104-109.
18. Kielkopf, C. L., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B., and Rees, D. C. (1998) A structural basis for recognition of A.T and T.A base pairs in the minor groove of B-DNA, *Science 282*, 111-115.
19. Hawkins, C. A., Clairac, R. P. d., Dominey, R. N., Baird, E. E., White, S., Dervan, P. B., and Wemmer, D. E. (2000) Controlling binding orientation in hairpin polyamide DNA complexes, *J. Am. Chem. Soc. 122*, 5235-5243.
20. Lacy, E. R., Le, N. M., Price, C. A., Lee, M., and Wilson, W. D. (2002) Influence of a terminal formamido group on the sequence recognition of DNA by polyamides, *J. Am. Chem. Soc. 124*, 2153-2163.
21. O'Hare, C. C., Mack, D., Tandon, M., Sharma, S. K., Lown, J. W., Kopka, M. L., Dickerson, R. E., and Hartley, J. A. (2002) DNA sequence recognition in the minor groove by crosslinked polyamides: The effect of N-terminal head group and linker length on binding affinity and specificity, *Proc Natl Acad Sci U S A 99*, 72-77.
22. Chenoweth, D. M., and Dervan, P. B. (2009) Allosteric modulation of DNA by small molecules., *Proc. Natl. Acad. Sci. 106*, 13175-13179.
23. Mrksich, M., Parks, M. E., and Dervan, P. B. (1994) Hairpin peptide motif: A new class of oligopeptides for sequence specific recognition in the minor groove of double helical DNA, *J. Am. Chem. Soc. 116*, 7983-7988.
24. Mrksich, M., and Dervan, P. B. (1993) Enhanced sequence specific recognition in the minor groove of DNA by covalent peptide dimers: bis(pyridine-2-carboxamideonetropsin)$(CH_2)_{3\text{-}6}$, *J. Am. Chem. Soc. 115*, 9892-9899.
25. Gottesfeld, J. M., Neely, L., Trauger, J. W., Baird, E. E., and Dervan, P. B. (1997) Regulation of gene expression by small molecules, *Nature 387*, 202-205.
26. Le, N. M., Sielaff, A., Cooper, A. J., Mackay, H., Brown, T., Kotecha, M., O'Hare, C., Hochhauser, D., Lee, M., and Hartley, J. A. (2006) Binding of f-PIP, a pyrrole- and imidazole-containing triamide, to the inverted CCATT box-2 of the topoisomerase IIα promoter and modulation of gene expression in cells *Bioorg Med Chem Letters 16*, 6161-6164.
27. Fortune, J. M., Velea, L., Graves, D. E., Utsugi, T., Yamada, Y., and Osheroff, N. (1999) DNA topoisomerases as targets for the anticancer drug TAS-103: DNA interactions and topoisomerase catalytic inhibition, *Biochemistry 38*, 15580-15586.
28. Fortune, J. M., and Osheroff, N. (2000) Topoisomerase II as a target for anticancer drugs: When enzymes stop being nice, *Prog. Nucleic Acid Res. Mol. Biol. 64*, 221-253.
29. Isaacs, R. J., Harris, A. L., and Hickson, I. D. (1996) Regulation of the human topoisomerase IIα gene promoter in confluence-arrested cells, *J. Biol. Chem. 271*, 16741–16747.

30. Tolner, B., Hartley, J. A., and Hochhauser, D. (2001) Transcriptional regulation of topoisomerase II alpha at confluence and pharmacological modulation of expression by bis-benzimidazole drugs, *D. Mol. Pharmacol. 59*, 699-706.

CHAPTER 2:

DNA RECOGNITION: DESIGN, SYNTHESIS AND BIOPHYSICAL CHARACTERISTICS OF PYRROLE (H) BASED POLYAMIDES

The work presented in this chapter is based on the published paper "*DNA Recognition: Design, Synthesis and Biophysical Characteristics of Pyrrole(H) based Polyamides."* from *Medicinal Chemistry*, **2010**, *6*, 150-158. Synthetic methods and write ups are attributed to Dr. Moses Lee of Hope College and various members of the Lee Research Group. Dr. John Hartley from the UK Drug-DNA Interactions Research Group is credited with the DNase I footprinting work and writings. My contribution to this paper was the SPR experiments, data analyses and writings, along with Dr. Manoj Munde and Dr. W. David Wilson.

DNA Recognition: Design, Synthesis and Biophysical Characteristics of Pyrrole(H) based Polyamides.

Sameer Chavda[a], Keith Mulder[a], Toni Brown[a], Hilary Mackay[a], Balaji Babu[a], Laura Westrate[a], Amanda Ferguson[a], Shicai Lin[c], Konstantinos Kiakos[c], Joseph P. Ramos[b], Manoj Munde[b], W. David Wilson[b], John A. Hartley[c] and Moses Lee[a,]

[a]*Department of Chemistry and the Division of Natural and Applied Sciences, Hope College 49423, USA*

[b]*Department of Chemistry, Georgia State University Atlanta, GA, 30303, USA*

[c]*Cancer Research, UK Drug-DNA Interactions Research Group, UCL Cancer Institute, Paul O' Gorman Building, 72 Huntley Street, London WCIE 6BT, UK*

2.1 Abstract

N-Methyl imidazole (Im) and *N*-methyl pyrrole (Py)-containing polyamides that can form stacked dimers can be programmed to target specific DNA sequences in the minor groove of DNA and control gene expression. Polyamides are being investigated as potential medicinal agents for treating diseases including cancer. The naturally occurring polyamide distamycin binds as a dimer in the minor groove of DNA and recognizes sequences rich in A/T and T/A base pairs indiscriminately. Synthetic analogs of distamycin that incorporate N-methylimidazole into the heterocyclic core have been shown to bind to G/C rich sequences with a high degree of specificity. The purpose of this study is to investigate the behavior of polyamides containing the 2, 5-linked N-methylpyrrole-2-carboxamide or pyrrole(H) [Py(H)] moiety upon binding to DNA. The synthesis and biophysical characteristics of two polyamides PyPyPyPy(H) **2** and ImPyPyPy(H) **3** designed to test the binding preference of a Py/Pyrrole(H) pairing [Py/Py(H)] and a [Im/Py(H)] is described. Studies utilizing circular dichroism, thermal denaturation (ΔT_M), biosensor-surface plasmon resonance and DNase I footprinting show that an [Im/Py(H), **3**] pairing prefers a G/C or C/G pairing whilst a [Py/Py(H), **2**] pairing tolerates A/T or T/A base pairs and avoids a G/C base pair.

Keywords: polyamides, pyrrole(H), DNA, sequence specificity, minor groove binder

2.2 Introduction

The development of polyamide analogs of distamycin **1** (Figure 2.1) that can target specific DNA sequences is a fascinating and well documented area of research. Such compounds exhibit potential for use as gene regulatory agents by inhibiting the binding of native transcription factors at the target promoter site, and are being investigated as potential medicinal agents for treating diseases including cancer [1]. The pyrrole (Py) heterocyclic moieties in distamycin stack as a 2:1 antiparallel dimer in the DNA minor groove and can bind indiscriminately to an A/T or a T/A base pair [2a-d]. Alternatively, an imidazole (Im) heterocycle moiety can be paired opposite a pyrrole to selectively target a G/C base pair [2d]. In addition, 3-hydroxy-1H-pyrrole (Hp) has been paired opposite Py and shown to selectively target T/A base pairs but with a significant loss in binding affinity. This could possibly attribute to a handful of factors ranging from sterics to chemical instability [3a-c]. The bulky hydroxyl group on an Hp containing polyamide (which forms an additional hydrogen bond to the O-2 group of thymine), makes the 2:1 complex formed in the minor groove sterically encumbered leading to unfavorable binding. Accordingly the additional hydrogen bonding to thymine O-2 could be critical for distinguishing A/T from T/A. We envisaged that a polyamide containing a reversed conformation Py moiety [Py(H) or 2,5-linked N-methylpyrrole-2-carboxamide] stacked over a polyamide containing Py [Py(H)/Py] may recognize a T/A base pair. In addition, it has been reported that the N-H group on this Py(H) moiety should form a more favorable van der Waals distance with thymine thus resulting in enhanced binding affinity [4]. In an attempt to test this hypothesis, we chose to synthesize two polyamides PyPyPyPy(H)-*N*-dimethyl-3-aminopropylamine **2** and ImPyPyPy(H)-*N*-dimethyl-3-aminopropylamine **3** (Figure 1). We chose to incorporate only one Py(H) moiety into each of these polyamides as previous studies have

shown that incorporation of multiple 2,5-linked [3a] or 2,4-linked Py(H) [4a] units may result in a loss in sequence specificity, thus allowing for an increased tolerance of G/C base recognition [4a]. Even though polyamides bind DNA effectively through the 2:1 stacked motif, they are also known to bind in a 1:1 fashion within the minor groove of DNA [4b].

2.3 Results and Discussion

2.3.1 Polyamide and DNA sequence design

Two tetraheterocyclic polyamides (tetraamides) **2** and **3** were synthesized to study the binding preference for stacked Py(H)/heterocycle pairings (Figure 2.2). These tetraamides contain one of two different terminal pairing motifs: –ImPy– and –PyPy–, which are the best pyrrole and imidazole containing terminal pairings in terms of binding affinity to Watson and Crick sequences [3d]. Two of the DNA sequences tested for these studies (AT_10 and TA_10) were designed such that the specificity of a Py/Py-H for a TA base pair could be assessed relative to that for an AT base pair. In addition, we were also interested to ascertain whether a stacked Py-H/Im would exhibit any selectivity for CG or GC base pairs to potentially uncover any new pairing rules. In addition, two additional sequences CG_10 and GC_10 (Figure 2) were included in the study. The binding of **2** and **3** against the 10 base pair sequences ACGCGT (Figure 2) was also undertaken. This sequence is of significant interest for drug design and discovery due to its G/C richness and presence in the core sequence of the Mlul cell-cycle box (MCB) transcriptional element indirectly implicated in the development of various cancers [5, 6, 7].

2.3.2 Synthesis

The synthetic approach towards obtaining polyamides **2** and **3** is outlined in Scheme 2.1. Nitration of **4** furnished 4-nitro-2-trichloroacetylpyrrole **5** in good yield, which in turn, was coupled with commercially available *N*,*N*-dimethyl-3-aminopropylamine in the presence triethylamine to furnish the monoheterocyclic component **6** [8]. This was followed by subsequent reduction of **6** using standard palladium catalyzed hydrogenation [9] to the resulting amino derivative (not shown) followed by immediate coupling to 4-nitro-1-methylpyrrole-2-carbonyl chloride **7** (formed from the corresponding carboxylic acid). The resulting diamide (O_2N-PyPy(H)-*N*-dimethyl-3-aminopropylamine, **8**), formed in 22% yield was, in turn, coupled to another molecule of 4-nitro-1-methylpyrrole-2-carbonyl chloride **7** using the protocol described, to furnish the corresponding triamide (O_2N-PyPyPy(H)-tail **9** in 39% yield (Scheme 2.1). At this point, we attempted to couple the relevant terminal heterocycle by forming the acid chloride of the relevant commercially available carboxylic acid (**10**, Scheme 2.1) and coupling this with the reduced amine form (Schotten-Baumann coupling) of **9** (formed by palladium catalyzed reduction in methanol) to give the final tetraamide **2** in 28% yield (Scheme 2.1). However, this did not seem to be the case for the formation of the tetraamide **3** due to low yields. An alternative set of coupling conditions utilizing the lithium salt **11** of imidazole-2-carboxylic acid [10] with benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBop) and *N*,*N*-diisopropyl-*N*-ethylamine (DiPEA) in dry DCM (Scheme 2.1), furnished tetraamide **3** in an improved yield. In general, one of the reasons for the low yields of these tetraamides and their tri- and diamide precursors, can be attributed to the instability of the amine intermediates [11] as well as the reactivity of the pyrrole and imidazole acid chlorides.

2.3.3 Thermal melting studies

Thermal melting studies were conducted to determine whether polyamides **2** and **3** stabilize duplex DNA at elevated temperatures. The data shown in Table 1 indicates that polyamide **2** can only produce a ΔT_M value of 3°C for the AT_10 sequence. TA_10, CG_10 and GC_10 did not increase the melting temperature in the presence of tetraamide **2**. On the other hand polyamide **3** exhibits some stabilization on sequences AT_10, TA_10 and GC_10 giving ΔT_M values of 2, 2 and 3°C, respectively but did not induce a ΔT_M for CG_10.

2.3.4 Circular dichroism studies

Circular dichroism studies (the ability of chiral molecules to absorb circularly polarized light: the DNA/polyamide complex is chiral due to the DNA's inherent chirality) were performed on PyPyPyPy(H)-*N*-dimethyl-3-aminopropylamine **2** and **3** ImPyPyPy(H)-*N*-dimethyl-3-aminopropylamine on the same DNA sequences used in the thermal denaturation studies (Figure 3). The results shown in Figure 2.3 for **3** confirm that it binds within the floor of the minor groove of the DNAs tested due to the appearance of an induced band at 330 nm [12]. The results also indicate that polyamide **2** gave weaker induced CD bands than **3** when titrated into solutions of the DNAs mentioned. Except for AT_10, the results were in agreement with the thermal denaturation data shown in Table 2.1. However, as depicted in Figure 3 polyamide **2** produced clear DNA induced ligand bands when titrated into solutions of AT_10 and TA_10 suggesting a limited preference of Py/Py(H) for an AT or TA base pair. The most impressive result was achieved when polyamide **3** was titrated into a solution of GC_10 with near-saturation being achieved after addition of two equivalents of **3**, further corroborated by a ΔT_M value of 3 °C. In addition, the appearance of a clear isodichroic point provides evidence that binding is

occurring via a single motif characteristic of a typical DNA minor groove binder. It is interesting to note that, though a ΔT_M value of 0°C was observed for polyamide **3** when titrated into a solution of CG_10 (Table 2.1), a significant amount of binding is observed in the CD study (Figure 3) in turn, suggesting binding of **3** to this DNA sequence is possibly thermodynamically less favorable than when binding to GC_10.

2.3.5 DNase I footprinting

DNase I footprinting experiments were performed on **2** and **3** using a 131 bp 5`-[$^{32}$P]-radiolabelled DNA fragment containing the following sequences: (a) 5`-AAATTT-`3; AT_10 (b) 5`-ATATAT-`3; TA_10 (c) 5`-ACATGT-`3; CG_10 (d) 5`-AGATCT-`3; GC_10 and (d) 5`-AGAGCT-`3. Autoradiogram A (Figure 2.4) shows that polyamide **2** exhibits good sequence specificity mostly for TA_10 DNA. A clear footprint appears at 20 μM, which is consistent with CD and ΔT_M measurements (Table 2.1 and Figure 2.3). Polyamide **2** also displays negligible binding to GC_10 as shown by the appearance of a weak footprint at 100 μM. In comparison, **3** seems to show better binding affinity towards the sequences tested but poorer sequence selectivity. Autoradiogram B (Figure 2.4) shows that for polyamide **3**, footprints at 5'-ATATAT-3', 5`-AGATCT-`3 and 5`-ACATGT-`3 are evident at 6 μM. This is corroborated by the thermal melting and CD data (Figure 2.3 and Table 2.1). In addition, enhanced cleavage appears to occur at the locations close to the AT_10 site. Enhanced cleavage sites have been reported for binding of polyamides to DNA [13].

2.3.6 Surface Plasmon Resonance

To obtain a more accurate measure of binding affinity, selectivity and to probe the stoichiometry of binding, SPR-biosensor experiments were performed. As described in the Methods Section, the biotin-labeled GC_10, AT_10 and TA_10 DNA hairpin duplexes were immobilized by streptavidin capture in three of the four flow cells. The fourth was left blank and used as a control. Compounds **2** and **3** were then injected over the DNA and blank flow cells at a range of concentrations. From the sensorgrams in Figure 2.5, it can be observed that **3** shows good selectivity for GC_10. The kinetics are relatively fast and all of the sensorgrams reach a steady state plateau below 200 s. Binding constants were determined by plotting the steady state RU values, determined by averaging over a 10 s time period in the steady state region, versus the compound concentration and fitting with either a one or two site binding model (Methods Section). Compound **3** binds strongly to GC_10 ($K = 2 \times 10^6$ M^{-1}, Table 2.1) with 2:1 stoichiometry but shows weaker binding towards AT_10 by 10-fold ($K = 2 \times 10^5$ M^{-1}) with the data best fitting a 1:1 stoichiometry. Very weak binding was observed for **3** with TA_10 and the stoichiometry could not be determined. Binding constants were determined by the same procedure for the pyrrole compound, **2**. It also has fast kinetics and shows selectivity for AT base pair sequences. Compound **2** shows significantly stronger binding to AT_10 than to TA_10 with the best fit to a 1:1 binding model in both cases (Table 2.1) and shows very weak binding to GC_10.

2.4 Conclusion

Overall, this study has shown that, polyamide **2** containing the Py/Py(H) pairing, displays weaker binding than expected to its cognate sequences. The results do show sequence selectivity (towards AT_10), in agreement with its predicted specificity, but with a lack of marked binding affinity. However, it prefers to avoid a G/C base pair, as ascertained by SPR, DNase I footprinting and thermal melting techniques. On the other hand, **3**, which contains an Im/Py(H) pairing, prefers a G/C or C/G base pair. For these reasons, we believe that additional polyamides related to **2** and **3** are worthy of further investigation.

2.4.1 Experimental

Solvents and organic reagents were purchased from Aldrich or Fisher, and in most cases were used without further purification. DCM (P_2O_5), and DMF (BaO) were distilled prior to use. Melting points (mp) were performed using a Mel-temp instrument and are uncorrected. Infrared (IR) spectra were recorded using a FT-IR instrument as films on NaCl discs. ^{1}H-NMR spectra were obtained using a Varian Unity Inova 400 instrument. Chemical shifts (δ) are reported at 20 °C in parts per million (ppm) downfield using tetramethylsilane (Me_4Si) as an internal standard. High-resolution mass spectra (HRMS) and low-resolution mass spectra (LRMS) were provided by the Mass Spectrometry Laboratory, University of South Carolina, Columbia. Reaction progress was assessed by thin-layer chromatography (TLC) using Merck silica gel (60 F_{254}) on aluminum plates unless otherwise stated. Visualization was achieved with UV light at 254 nm and/or 366 nm, I_2 vapor staining and ninhydrin spray.

2.4.2 Synthesis

2-Trichloroacetylpyrrole 4 and 5-nitro-2-trichloroacetylpyrrole 5

These precursor compounds were synthesized according to a literature protocol and were spectroscopically identical to those reported [3].

O_2N-Py(H) *N,N*-dimethyl-3-aminopropylamine 6

O_2N-Py(H)-CCl_3 (500 mg, 1.95 mmol) was dissolved in dry CH_2Cl_2 (60 mL) and triethylamine (0.41 mL, 2.93 mmol), *N*,*N*-dimethyl-3-aminopropylamine (0.37 mL, 2.93 mmol) was added and the reaction mixture was heated at reflux for ~18 h. The reaction was quenched with water (50 ml) and the resulting aqueous basified to pH 10 was and the organic layer was collected. The aqueous was extracted with $CHCl_2$ (2 × 40 mL). The combined organic layers were dried over Na_2SO_4 and evaporated to dryness. The crude product was purified using flash chromatography on silica gel eluting with to yield **6** as an orange/yellow solid (300 mg, 68%); R_F 0.46 (silica gel, 40:39.5:0.5 MeOH:$CHCl_3$:NH_4OH); mp 154-160 ºC; υ (cm^{-1}) 1692 (C=O); δ_H (400 MHz; $CDCl_3$) 7.65 (1 H, d, *J* 1.6 Hz); 7.15 (1 H, d, *J* 1.6), 3.37 (2 H, t, *J* 6.8), 2.49 (2 H, t, *J* 6.8); 2.34 (6 H, s) and 1.76 (2 H, quintet, *J* 6.8); LRMS: (direct probe) *m/z* (rel. intensity) 240 ([M^+], 5%), 139 (4%), 93 (4%), 72 (5%).

O_2N-Py-Py(H) *N,N*-dimethyl-3-aminopropylamine 8

A round bottom flask purged with argon, containing O_2N-Py(H)-*N*-dimethyl-3-aminopropylamine, **6** (186 mg, 0.82 mmol), 5% Pd/C (93 mg, 50% by wt.) cold MeOH (60 ml) was cautiously added and the resulting solution was stirred under H_2 for 12 h. The Pd/C catalyst was filtered over a pad of celite and the filtrate co-evaporated with $CHCl_2$ (3 × 20 mL) to give a

yellow/brown oil which was stored under high vacuum out of light. NO_2-Py-COOH (0.168 g, 0.987 mmol) was dissolved in $SOCl_2$ (3 mL) and refluxed for 15 minutes. The resulting solution was co-evaporated with $CHCl_2$ (3 × 15 mL) to give the acid chloride **7** as an oil which was diluted with dry CH_2Cl_2 (5 mL) and added dropwise to a stirring solution of the reduced form of **6** and dry triethylamine (0.14 ml, 0.99 mmol) in dry $CHCl_2$ (10 mL) cold (0 ºC, ice/water). The reaction mixture was stirred at room temperature overnight. A basic aqueous work-up was performed and the organic layers collected, dried over (Na_2SO_4) and evaporated to dryness. The crude product was purified using flash-column chromatography to yield **8** as a yellow solid (67 mg, 22%), mp. decomposition at 200 ºC; R_F 0.20 (silica gel, 80:19.5:0.5 % v/v; $CHCl_3$:MeOH:NH_4OH); υ (CH_2Cl_2, cm^{-1}) 1640 (C=O) 1636 (C=O); δ_H (400 MHz, $CDCl_3$) 10.08 (1 H, s (br), 9.24 (1 H, s (br)), 7.74 (1 H, s (br)), 7.59 (1 H, d, *J* 1.6), 7.37 (1 H, s), 7.36 (1 H, d, *J* 1.6), 6.70 (1 H, s), 4.02 (3 H, s), 3.42 (2 H, t, *J* 6.4), 2.44 (2 H, t, *J* 6.4), 2.30 (6 H, s) and 1.76 (2 H, quintet, *J* 6.4); MS: (ES^+) *m/z* (rel. intensity) 363 ([M+H] 100%), 255 (5%).

O_2N-Py-Py-Py(H) *N,N*-dimethyl-3-aminopropylamine 9

In the same way as for **8**, O_2N-Py-Py(H)-*N*-dimethyl-3-aminopropylamine **8** (67 mg, 0.184 mmol), 5% Pd/C (35 mg, 50% by wt.) and cold MeOH (40 mL) gave the corresponding amine as a yellow oil. Also in the same way as **8**, NO_2-Py-COCl **7** (from NO_2-Py-COOH, 37 mg, 0.22 mmol and 3) and dry triethylamine (0.03 ml, 0.22 mmol) gave crude **9** which was purified using flash chromatography eluting in to yield **9** as a yellow/brown solid (35 mg, 39%); mp decomposition at 200 ºC: R_F 0.3 (silica gel, 60:39.5:0.5 $CHCl_3$:MeOH:NH_4OH); υ (CH_2Cl_2, cm^{-1}) 1640 (C=O) 1636 (C=O), 1638 (C=O); δ_H (400 MHz, $CDCl_3$) 7.57 (1 H, d, *J* 2.0); 7.35 (1 H, d, *J* 1.6), 7.33 (1 H, d, *J* 1.6); 7.24 (1 H, d, *J* 1.6), 6.77 (1 H, *J* 1.6), 6.72 (1 H, d, *J* 1.6); 3.99 (3

H, s), 3.90 (3 H, s), 3.39 (2 H, t, *J* 6.4), 2.63 (2 H, t, *J* 7.2); 2.45 (6 H, s) and 1.83 (2 H, quintet, *J* 7.2); LRMS: (ES^+) *m/z* (rel. intensity) 485 ([M+H], 100%), 377 (20%).

Py-Py-Py-Py(H) *N,N*-dimethyl-3-aminopropylamine 2

In the same way as for the method stated in for the reduction of **8** and coupling of the resultant amine, O_2N-Py-Py-Py(H)-*N*-dimethyl-3-aminopropylamine **9** (75 mg, 0.155 mmol), 5% Pd/C (40 mg, 50% by wt) and cold MeOH (8 mL) gave the resultant amine as a orange/yellow oil. The amine from **11**, Py-COCl **10** (from 39 mg, 0.309 mmol) pyrrole-2-carboxylic acid and dry triethylamine (0.04 mL, 0.30 mmol) gave a crude product was purified using flash chromatography to yield **2** as an orange solid (25 mg, 28%), mp decomposition at 210 ºC; R_F 0.30 (silica gel, 60:39.5:0.5 $CHCl_3$/MeOH/NH_4OH); 1640 (C=O) 1636 (C=O), 1638 (C=O) and 1637 (C=O); δ_H (400 MHz, $CDCl_3$) 7.77 (1 H, d, *J* 2.0); 7.52 (1 H, d, *J* 1.6), 7.33 (1 H, d, *J* 1.6); 7.23 (1 H, d, *J* 1.6), 7.12 (1 H, s), 6.77 (2 H, d, *J* 1.6), 6.10 (1 H, d, *J* =1.6), 6.09 (1 H, d, *J* 1.6), 5.55 (2 H, 2 × s (br)), 5.15 (2 H, 2 × s (br)), 3.93 (3 H, s), 3.90 (3 H, s), 3.89 (3 H, s), 3.35 (2 H, t, *J* 7.2), 2.63 (2 H, t, *J* 7.2); 2.55 (6 H, s) and 1.85 (2 H, quintet, *J* 7.2); LRMS: 562 ([M+H], 100%), 282 (80%), 145 (90%); HRMS: $C_{28}H_{36}N_9O_4$ requires 562.2890; found 562.2901.

Im-Py-Py-Py(H) *N,N*-dimethyl-3-aminopropylamine 3

The reduction of **9** was achieved by following the same protocol as stated in for **8** using the following: O_2N-Py-Py-Py(H)-tail **9** (75 mg, 0.155 mmol), 5%Pd/C (40 mg, 50% by wt.) and cold MeOH (65 mL). The lithiated salt of 1 methylimidazole-2-carboxylic acid **11**, from 39 mg, 0.309 mmol 1-methylimidazole-2 carboxylic acid) and PyBOP (39 mg, 0.309 mmol) were dissolved in dry DCM (10 mL) under argon. DiPEA (0.084 mL, 0.33 mmol) was added to the solution and the reaction mixture was stirred at 50 ºC for 3 days. A basic aqueous work-up was performed

(similar to that of described previously and the organic layer collected, dried over (Na_2SO_4) and evaporated to dryness. The crude product was purified using flash-column chromatography to yield **3** as a yellow solid (15 mg, 17%), mp. decomposition at 210 ºC; R_F 0.40 (silica gel, 60:39.5:0.5 $CHCl_3$:MeOH:NH_4OH); υ (neat, cm^{-1}) 1640 (C=O) 1636 (C=O), 1638 (C=O) and 1635 (C=O); δ_H (400 MHz, $CDCl_3$); 10.32 (1 H, s (br)), 9.45 (1 H, s (br)), 8.48 (1 H, s (br), 8.27 (1 H, s (br)) 7.89 (1 H, s (br)), 7.34 (1 H, d, *J* 2.4), 7.21 (1 H, d, *J* 2.4), 7.00 (1 H, t, *J* 4.8), 6.94 (1 H, s), 6.79 (1 H, s), 6.68 (1 H, d, *J* 6.5), 6.65 (1 H, d, *J* 6.5), 4.03 (3 H, s), 3.88 (3 H, s,), 3.38 (3 H, s), 3.40 (2 H, t, *J* 7.2), 2.45 (2 H, t, *J* 7.2), 2.21 (6 H, 2 × s) and 1.71-1.63 (2 H, quintet, *J* 7.2); LRMS: (ES+) *m/z* (rel. intensity) 563 ([M+H], 15%), 297 (45%), 282 (100%), 145 (20%); HRMS, $C_{27}H_{35}N_{10}O_4$ requires 563.2842; found $C_{27}H_{35}N_{10}O_4$ 563.2847.

2.4.3 Biophysical

All buffer reagents were obtained from Sigma-Aldrich or Fisher, and used without further purification unless stated otherwise. The synthetic oligonucleotides were obtained from Operon (Huntsville, AL): **ACGCGT** (5'-ACGCGT-3'): 5'-GA ACGCGT CG CTCT CGACGCGTTC-3'; **A_3T_3_10** (5'-A_3T_3-3'): 5'-CG AAATTT CC CTCT GG AAATTT CG-3'; **AT_10** (5'-AT-3'): 5'-CC AAATTT CC CTCTGG AAATTT GG; **TA_10** (5'-TA-3'): 5'-CC ATATAT CC CTCTGG ATATAT GG; **CG_10** (5'-CG-3'); 5'-CC ACATGT CC CTCTGG ACATGT GG and **GC_10** (5'-GC-3'); 5'-CC AGATCT CC CTCTGG AGATCT GG were used in the thermal denaturation and circular dichroism experiments without further purification. DNA used in the footprinting and SPR experiments are described below. CD data was obtained using an Olis DSM20 instrument. ΔT_M data were obtained using a Varian-Cary 100 Biomelt UV–visible spectrophotometer, equipped with a Peltier temperature controller and a precision cuvette-mounted temperature probe. Phosphate buffer: $PO_4$0 (10 mM sodium phosphate, 1mM EDTA,

pH 6.4) was used for CD and T_m experiments. All data were analyzed using KaleidaGraph (Snergy Software, Reading PA), unless otherwise stated. Ligand stock solutions were prepared in double distilled water at a concentration of 0.5 mM, unless otherwise specified.

2.4.4 Circular dichroism (CD)

CD studies were performed using previously reported procedures [14b,f] and were conducted at ambient temperature in a 1-mm path length quartz cell using PO_4 buffer. Buffer and stock DNA were added to the cuvette to give a final DNA concentration of 9 μM. Each polyamide (in 500μL in double distilled) H_2O was titrated in 1 molar equivalents into the relevant DNA (160 μL of 9 μM DNA) until saturation was achieved. Each run was performed over 400-220 nm. The CD response at the λ_{max} of the induced peak was plotted against the mole ratio of ligand:DNA.

2.4.5 Thermal denaturation (ΔT_M)

Thermal denaturation studies were performed using published procedures [3d]. Experiments for polyamides **2** and **3** were performed at a concentration of 3 μM ligand and 1 μM DNA. All experiments were run in PO_4 buffer. Oligonucleotide samples were reannealed prior to denaturation studies by heating at 70 ºC for 1 min then cooling to RT. Heating runs were typically performed between 25 and 95 ºC, with a heating rate of 0.5 ºC min^{-1}, while continuously monitoring the absorbance at 260 nm (digitally sampled at 200 ms intervals). All melts were performed in 10-mm path length quartz cells. T_m values were determined as the maximum of the first derivative.

2.4.6 Surface plasmon resonance (SPR)

Buffers: a 0.01 M cacodylic acid (CCA) solution pH 6.25 contained 0.001 M EDTA (disodium ethylenediamine tetraacetate), 0.1 M NaCl and 0.005% (by vol.) surfactant P20 which was (Biacore AB) used to reduce the nonspecific binding of polyamides to the fluidics and sensor chip surface. In addition, a 0.01 M HEPES (N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]) solution pH 7.4 containing 0.15 M NaCl, 0.003 M EDTA and 0.005% (by volume) surfactant P20 was used during the DNA immobilization process. All buffers were degassed and filtered prior to experiments. DNA sequences were obtained from Integrated DNA Technologies (San Diego, CA) with HPLC purification and were used without further purifying. The lyophilized 5' biotin-labeled DNA hairpin constructs were dissolved in the appropriate amount of DI H_2O to create 1.0 mM DNA stock solutions. Further dilutions were made using 0.01 M HEPES buffer during the DNA immobilization step. DNA concentrations were determined spectrophotometrically using extinction coefficients calculated for each individual DNA sequence by using the nearest neighboring for singled stranded DNA method. Stock solutions of **4** and **5** were prepared by dissolving the solid compound in the necessary amount of distilled H_2O to create a stock solution with a concentration of approximately 1.0×10^{-3} M. Stock solutions were kept frozen at 4 ˚C until experimental use to minimize any degradation of the compound. Samples for SPR experiments were prepared by a series of dilutions from the stock solution using a 0.01 M cacodylic acid buffer solution.

The surface plasmon resonance (SPR) experiments were conducted to analyze real time interactions between the polyamide compounds and specific DNA sequences. All SPR experiments were conducted using a four flow cell Biacore 2000 biosensor instrument (GE Life Sciences). DNA hairpin constructs labeled with biotin at the 5' end were immobilized onto a

streptavidin-coated sensor chip (BIAsensorSA) as previously reported [14]. The 5'-biotin-labeled DNA hairpins were immobilized on three of the four flow cells. The fourth was left blank and used as a control. All SPR experiments were performed at 25 ˚C and used 0.01 M CCA as the running buffer. The amount of DNA immobilized was approximately 400 response units (RU). This was achieved by continuously injecting ~ 20 µL of an approximately 50 nM DNA solution at a rate of 2 µL/min onto the sensor chip surface until a relative response of 400 units was reached. Binding data was obtained by injecting known concentrations and were analyzed with one or two site binding models as previously described [14].

2.4.7 DNase I Footprinting Studies

The designed DNA fragment was amplified by using PCR as follows. The forward primer 5'- GTCGGTTAGGAGAGCTCCACTTG-3' (4 pmol) was 5' ^{32}P end-labeled with [-^{32}P] ATP using T4 kinase (Invitrogen) following standard protocols. PCR amplification was performed in a reaction mixture of 50 µL containing the ^{32}P labeled primer, 4 pmol of reverse primer 5'-CTCCAGAAAGCCGGCACTCAG-3', 2 µL of 25 mM $MgCl_2$, 2.5 µL of 2.5 mM dNTPs (Promega, UK), 1U Flexi GoTaq polymerase (Promega,UK), the supplied 5x GoTaq buffer, and 0.125 pmol of each of the templates

5'ATGCTCCAGAAAGCCGGCACTCAGTCTACAAAAATTTCATCTTGATCATATATGTT CACAGAGCTCTCTCTAGATCTAGATCTAACTCTAGTACATGTCTTCAAGCAAGTGGA GCTCTCCTAACCGACTTT-3′ and

5′-AAAGTCGGTTAGGAGAGCTCCACTTGCTTGAAGACATGTACTAGAGTTA

GATCTAGATCTAGAGAGAGCTCTGTGAACATATATGATCAAGATGAAATTTTTGTAG ACTGAGTGCCGGCTTTCTGGAGCAT-3'.

PCR cycling conditions consisted of an initial denaturation step for 3 min at 95 °C, and 1 min at 94 °C, 1 min at 63 °C, and 1 min at 72 °C for a total of 35 cycles. The amplified fragment was isolated on a 2% agarose gel electrophoresis and purified with the Mermaid Kit (Q-biogene) according to the manufacturer's instructions.

DNase I digestions were conducted in a total volume of 8 µL. The labeled DNA fragment (2 µL, 200 counts s^{-1}) was incubated for 30min at room temperature in 4 µL TN binding buffer (10 mM Tris Base, 10 mM NaCl, pH 7) containing the required drug concentration. Cleavage by DNase I was initiated by addition of 2 µL of DNase I solution (20 mM NaCl, 2 mM $MgCl_2$, 2 mM $MnCl_2$, DNase I 0.02U, pH 8.0) and stopped after 3 min by snap freezing the samples on dry ice. The samples were subsequently lyophilized to dryness and resuspended in 5 µL of formamide loading dye (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% cyanol blue). Following heat denaturation for 5 min at 90^{o}C, the samples were loaded on a denaturing polyacrylamide (10%) gel (Sequagel, National Diagnostics, UK) containing urea (7.5 mM). Electrophoresis was carried out for 2 hours at 1650 V (~70W, 50°C) in 1X TBE buffer. The gel was then transferred onto Whatman 3MM and dried under vacuum at 80°C for 2 h. The gel was exposed overnight to medical X-Ray film (Super RX, Fuji) and developed on a Konica Medical Film Processor SRX-101A,

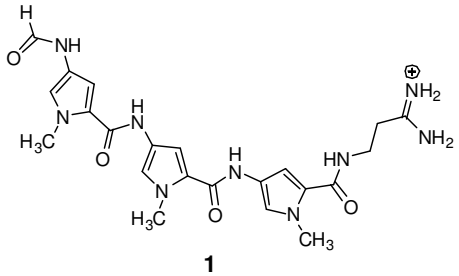
2.5 Acknowledgements

The authors thank the NSF (CHE 0550992 and 0809162) and Cancer Research UK (C2259/A9994) for support.

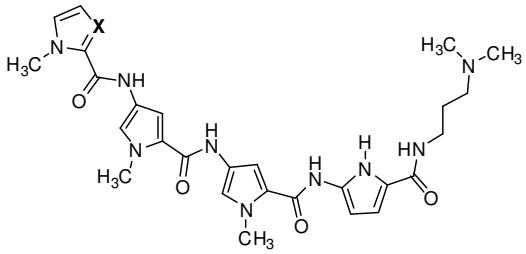
2.6 References and notes

1. (a) Melander, C; Burnett, R; Gottesfeld, J. M. Regulation of Gene Expression with Pyrrole-Imidazole Polyamides *J. Biotechnol.* **2004**, *112*, 195-220; (b) Dervan, P. B.; Edelson, B. S. Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284-299; (c) Reddy, B. S.; Sharma, S. K.; Lown, J. W. Recent developments in sequence selective minor groove DNA effectors. *Curr. Med. Chem.* **2001**, *8*, 475-508.
2. (a) Wade, W. S.; Mrksich, M; Dervan, P. B. Design of peptides that bind in the minor groove of DNA at 5'-(A,T)G(A,T)C(A,T)-3' sequences by a dimeric side by side motif. *J. Am. Chem. Soc.* **1992**, *114*, 8783-8794; (b) Pelton, J. G.; Wemmer, D. E. Binding modes of distamycin A with d(CGCAAATTTGCG)2. *J. Am. Chem. Soc.* **1990**, *112*, 1393-1399; (c) Wade, W. S.; Mrksich, M; Dervan, P. B. Binding Affinities of Synthetic Peptides, Pyridine-2-carboxamide-netropsin and 1-Methylimidazole-2-carboxamide-netropsin, that Form 2:1 Complexes in the Minor Groove of Double-Helical DNA. *Biochemistry* **1993**, *32*, 11385-11389; (d) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *89*, 1376-1399.
3. (a) Marques, M. A.; Doss, R. M.; Urbach, A. R.; Dervan, P.B. Toward an Understanding of the Chemical Etiology for DNA Minor-Groove Recognition by Polyamides. *Helv. Chim. Acta.* **2002**, *85*, 4485-4517; (b) Kielkopf, C. L.; Bremer, R. E.; White, S.; Szewcyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan; P. B. Structural Effects of DNA Sequence on T.A Recognition by Hydroxypyrrole/Pyrrole Pairs in the Minor Groove. *J. Mol. Biol.* **2000**, *295*, 557-567; (c) Kielkopf, C. L.; White, S.; Szewcyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P.B., Rees, D. C. A Structural Basis for Recognition of A… and T… Base Pairs in the Minor Groove of B-DNA. *Science*, *(Washington D.C.)* **1998**, *282*, 111-121; (d) Buchmueller, K. L.; Staples, A. M.; Howard, C. M.; Horrick, S. M.; Uthe, P. B.; Minh Le, N.; Cox, K. K.; Nguyen, B.; Pacheco, K. A.O.; Wilson, W. D.; C.; Lee, M. Extending the Language of DNA Molecular Recognition by Polyamides: Unexpected Influence of Imidazole and Pyrrole Arrangement on Binding Affinity and Specificity. *J. Am. Chem. Soc.*, **2005**, *127*, 742-750.
4. (a) Bremer, R. E.; Szewcyk, J. W.; Baird, E. E.; Dervan; P. B. Recognition of the DNA Minor Groove by Pyrrole-Imidazole Polyamides: Comparison of Desmethy- and N-Methylpyrrole. *Bioorg. Med. Chem.* **2000**, *8*, 1947-1955; (b) Urbach, A. R.; Dervan, P. B. Toward rules for 1:1 polyamide:DNA recognition. *Proc Natl Acad Sci U S A.* **2001**, *98*, 4343-4348.
5. Verma, R.; Patapoutian, A.; Gordon, C.B.; Campbell, J. L. Identification and purification of a factor that binds to the Mlu I cell cycle box of yeast DNA replication genes. *Proc. Natl. Acad.Sci. U.S.A.* **1991**, *88*, 7155-7159.
6. (a) Wu. X.; Lee, H. Human Dbf4/ASK promoter is activated through the Sp1 and *Mlu*I cell-cycle box (MCB) transcription elements. *Oncogene* **2002**, *21*, 7786-7796; (b) Yamada, M.; Sato, N.; Taniyama, C.; Ohtani, K.; Arai, K.; Masai, H. *J. Biol. Chem.* **2002**, *277*, 27668-27681.
7. Hess, G. F.; Drong, R. F.; Weiland, K. L.; Slightom, J. L.; Sclafani, R. A.; Hollingsworth, R. E. A human homolog of the yeast CDC7 gene is overexpressed in some tumors and transformed cell lines. *Gene* **1998**, *28*, 133-140.

8. Jaramillo, D.; Liu.; Aldrich-Wright, J.; Tor, Y. Synthesis of N-Methylpyrrole and N-Methylimidazole Amino Acids Suitable for Solid-Phase Synthesis. *J. Org. Chem.*, **2004**, 69, 8151-8153.
9. Buchmueller, K. L.; Bailey, S. L.; Matthews, D. A.; Taherbhai, Z. T.; Register, J. K.; Davis, Z. S.; Bruce, C. D.; O'Hare, C.; Hartley, J. A.; Lee, M. Physical and Structural basis for the Strong Interactions of the -ImPy- Central Pairing Motif in the Polyamide f-ImPyIm. *Biochemistry*, **2006**, *45*, 13551-13565.
10. Heckel, A.; Dervan, P. B. U-pin Polyamide Motif for Recognition of the DNA Minor Groove. *Chem. Eur. J.* **2003**, *9,* 3353-3366.
11. Yields for tri and diamides synthesized reported previously in the author's laboratory have typically ranged from 20-40% despite the development of specific handling techniques in an attempt to minimize the losses at each stage in the construction of each polyamide, see Brown, T.; Mackay, H.; Turlington, M.; Sutterfield, A.; Smith, T.; Sielaff, A.; Westrate, L.; Bruce, C.; Kluza, J.; O'Hare, C.; Nguyen, B.; Wilson, D.W.; Hartley, J. A.; Lee, M. Modifying the N-terminus of polyamides: PyImPyIm has improved sequence specificity over f-ImPyIm. *Bioorg. Med. Chem.* **2008**, *16*, 5266-5276.
12. (a) Lyng, R.; Rodger, A.; Nordén. B. The CD of ligand-DNA systems. 2. Poly(dA-dT) B-DNA. *Biopolymers* **1992**, *32*, 1201-1214; (b) Lyng, R.; Rodger, A.; Nordén. B. The CD of ligand-DNA systems. I. Poly(dG-dC) B-DNA. *Biopolymers* **1991**, *31*, 1709-1720.
13. (a) Lown, W. J.; Krowicki, K,; Bhat. U. G.; Skorobogarty, A.; Ward, B.; Dabrowiak, J. C. Molecular recognition oligopeptides and nucleic acids: novel imidazole-containing oligopeptides related to netropsin that exhibit altered DNA sequence specificity. *Biochemistry* **1986**, *25*, 7408-7416; (b) Lown, W. J.; Sondhi, S.M.; Ong, C. W. Skorobogarty, A.; Kishikawa, H.; Dabrowiak, J. C. Deoxyribonucleic acid cleavage specificity of a series of acridine and acodazole –iron porphyrins as functional bleomycin models. *Biochemistry* **1986**, *25*, 5111-5117.
14. (a) Lacy, E. R., Nguyen, B.; Minh Le, N.; Cox, K. K.; O' Hare, C.; Hartley, J. A.; Lee, M.; Wilson, D. Energetic basis for selective recognition of T*G mismatched in DNA by imidazole-rich polyamides. *Nucleic Acids Research*, **2004**, *32*, 2000-2007; (b) Lacy, E. R.; Minh Le, N.; Price, C. A.; Lee, M.; Wilson, D. Influence of the Terminal Formamido Group on the Sequence Recognition of DNA by polyamides. *J. Am. Chem. Soc.* **2002**, *124*, 2153-2163; (c) Nguyen, B.; Tanious, F.; Wilson. Biosensor Surface Plasmon Resonance: Quantitative Analysis of Small-Molecule Nucleic Acid Interactions. *Methods*. **2007**, *42*, 150-161; (d) Munde, M.; Ismail, M. A.; Arafa, R.; Peixoto, P.; Collar, C. J.; Liu, Y.; Hu, L.; David-Cordonnier, M-H.; Lansiaux, A.; Bailly, C.; Boykin, D. W.; Wilson, D. W.; Design of Minor Groove Binding Diamidines That Recognize GC Base Pair Sequences: A Dimeric Hinge Interaction Motif. *J. Am. Chem. Soc.* **2007**, *129*, 13732-13743; (e) Buchmueller, K.L.; Staples, A.M.; Howard, C.M.; Horick, S.M.; Uthe, P.B.; Le, N.M.; Cox, K.K.; Nguyen, B.; Pacheco, K.A.O.; Wilson, W.D.; Lee, M. *Nucleic Acids Research*, **2005**, *33*, 912-921; (f) Lacy, E. R., Nguyen, B.; Minh Le, N.; Cox, K. K.; O' Hare, C.; Hartley, J. A.; Lee, M.; Wilson, D. Recognition of T*G Mismatched Base Pairs in DNA by Stacked Imidazole-Containing Polyamides- Surface Plasmon Resonance and Circular Dichroism Studies. *Nucleic Acids Research*, **2002**, *30*, 1834-1841.



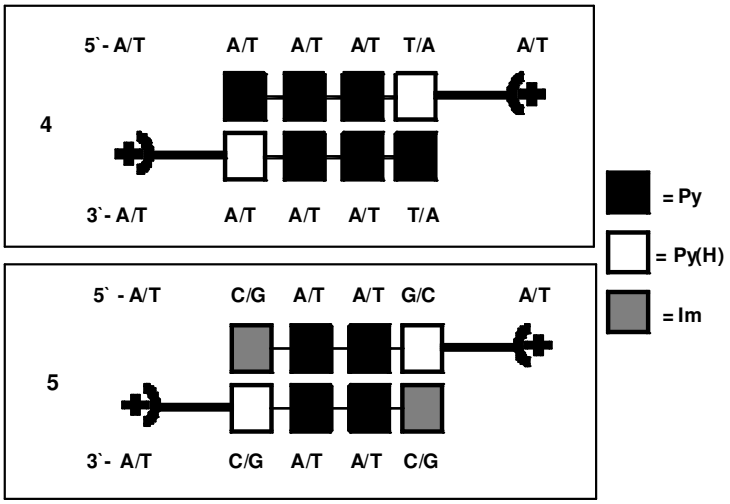
1



2; X=CH

3; X=N

Figure 2.1 Structures of Distamycin **1**, synthesized polyamides **2** and **3** designed to differentiate between A and T base pairs: PyPyPyPy(H)-*N,N*-dimethyl-3-aminopropylamine **2** and ImPyPyPy(H))-*N,N*-dimethyl-3-aminopropylamine **3**.



AT_10; 5'-AT-3': 5'-CG AAATTT CC CTCT GG AAATTT CG-3'

TA_10; 5'-TA-3': 5'-CC ATATAT CC CTCTGG ATATAT GG- 3'

CG_10; 5'-CG-3'; 5'-CC ACATGT CC CTCTGG ACATGT GG-3'

GC_10; 5'-GC-3'; 5'-CC AGATCT CC CTCTGG AGATCT GG-3'

ACGCGT; 5'-ACGCGT-3': 5'-GA ACGCGT CG CTCT CGACGCGTTC-3'

Figure 2.2 Binding of **2** and **3** to DNA in a 2:1 motif along with the 10 base pair DNA sequences to test any new pairing rules.

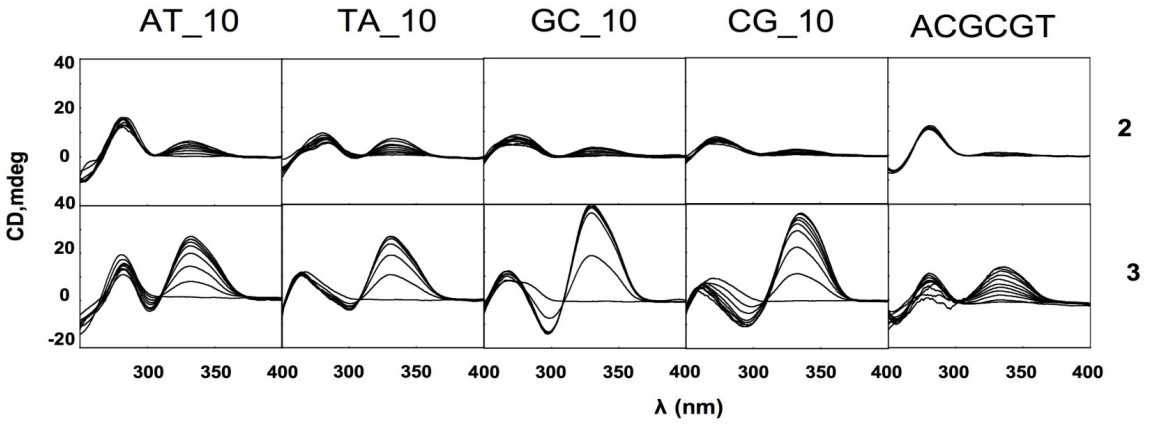


Figure 2.3 Circular dichroism data for **2** and **3**. Note that the horizontal axis represents wavelength (nm) whilst the CD in millidegrees (mdeg) is represented on the vertical axis. Both polyamides **2** and **3** (each in 500μL in double distilled H_2O) were titrated in 1 molar equivalents into the relevant DNA (160 μL of 9 μM DNA) until saturation was achieved.

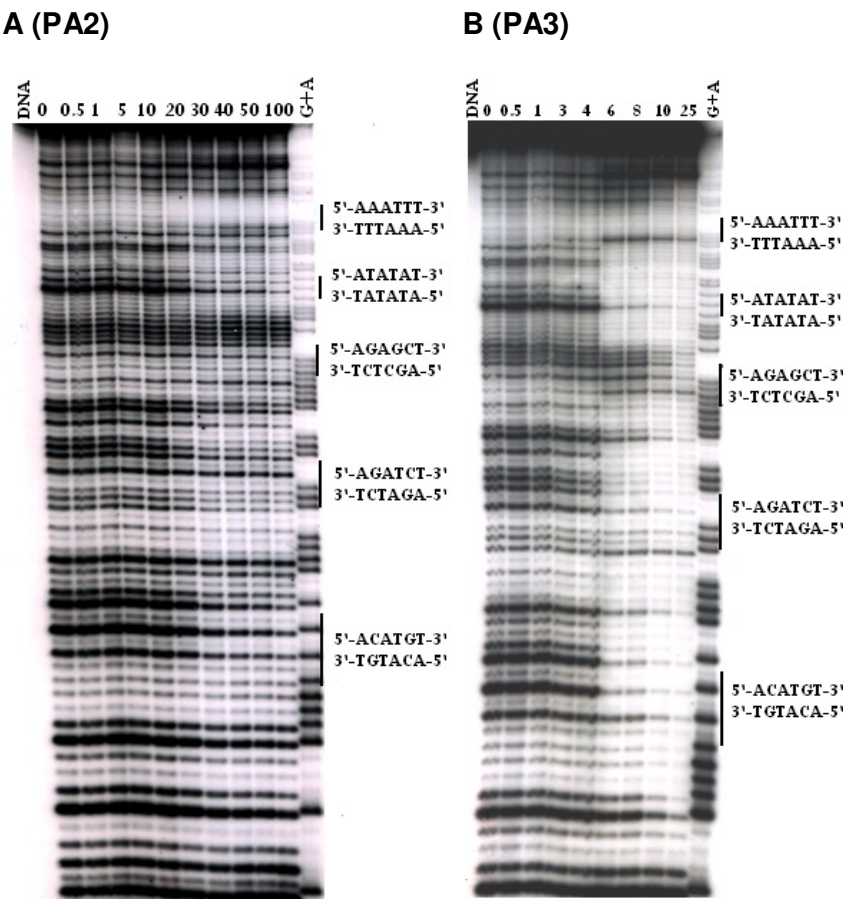


Figure 2.4 DNase I footprinting data of polyamides **2** (A) and **3** (B).

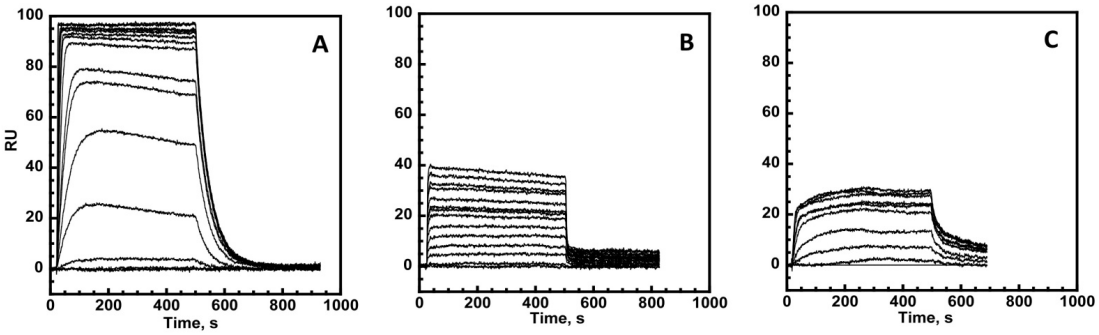
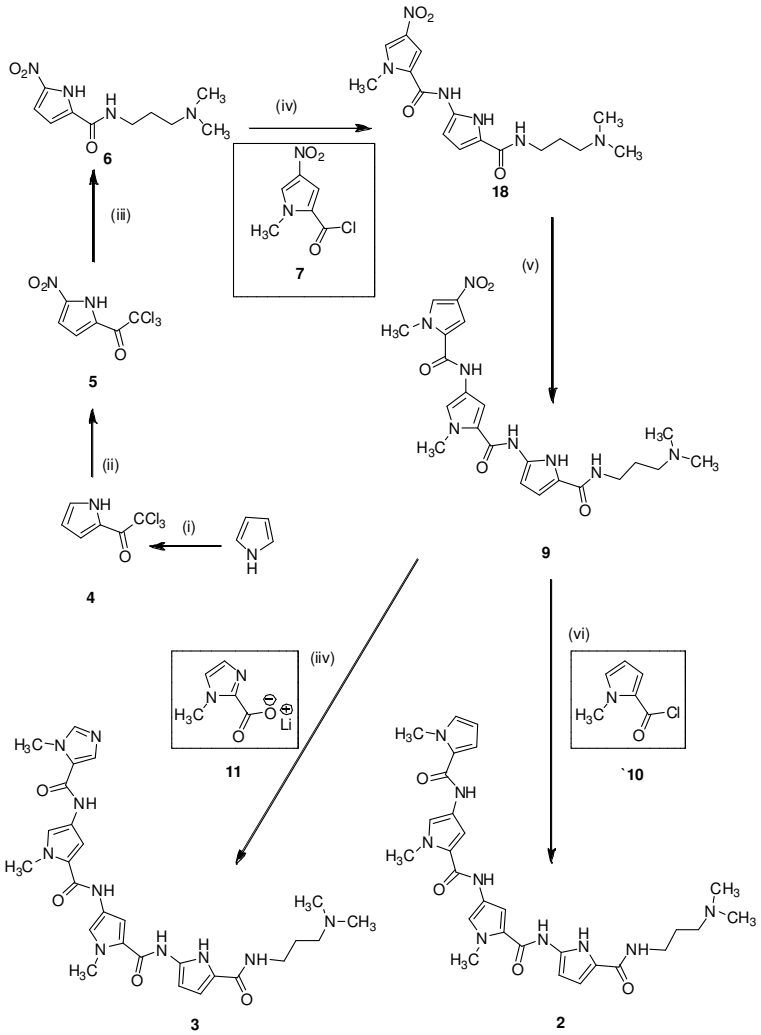


Figure 2.5 SPR sensorgrams of **3** with the sequence AGATCT, GC_10 (A), **3** with AT_10 (B), **2** with AT_10 (C). The ligand concentrations on the sensorgrams from the lowest to highest response are 0.1, 0.2, 0.3, 0.5, 0.7, 0.9, 1, 1.3, 1.5, 1.7, 2, 2.3, 2.5, 2.7, 3, 3.5, 4, 4.5 and 5 μM. All sensorgrams were plotted on the same scale and concentration range for comparison.



Scheme 2.1 Synthesis of **2** and **3**. Reagents and conditions: (i) $ClCOCCl_3$, CH_2Cl_2; (ii) Ac_2O; HNO_3; (iii) Et_3N, *N*,*N*-dimethyl-3-aminopropylamine, dry DCM, reflux, ~ 18 h; (iv) 5% Pd/C, cold MeOH, H_2, RT, ~ 12 h then Et_3N, dry DCM, acid chloride **7**; 0 °C-RT 12 hrs; (v) 5% Pd/C, cold MeOH, H_2, RT~ 12 h, ; dry Et_3N, dry DCM, reduced amine from **8**, 0-25 ºC, ~ 12 h. Synthesis of **2**: (vi) 5% Pd/C, cold MeOH, H_2, RT, ~ 18 h then **10**, dry Et_3N, dry DCM, 0 °C then warm to RT, stir 12 hrs. Synthesis of **5**; (iiv) 5% Pd/C, cold MeOH, H_2, RT, ~ 18 h then **11**, PyBOP, dry DCM, DiPEA, RT-50 ºC, 3 days.

Table 2.1 Thermal denaturation data for polyamides **2** and **3** and binding constants for polyamides **2** and **3** (as ascertained from SPR) for the sequences TA_10, GC_10 and A_3T_3_10.

	Sequence				
Polyamide	**AT_10**	**TA_10**	**GC_10**	**CG_10**	ACGCGT
(ΔT_M) PyPyPyPy(H)-*N*-dimethyl-3-aminopropylamine **2**	3ºC	0ºC	0ºC	0ºC	0ºC
K_a (polyamide **2**) (M^{-1})	1×10^6	$<1 \times 10^5$	$<<1 \times 10^5$	ND	ND
(ΔT_M) ImPyPyPy(H)- *N*-dimethyl-3-aminopropylamine **3**	1ºC	2ºC	3ºC	0ºC	0ºC
K_a (polyamide **3**) (M^{-1})	2×10^5	$<1 \times 10^5$	2×10^6	ND	ND

ND-not determined

CHAPTER 3:

SYNTHESIS AND DNA BINDING PROPERTIES OF 1-(3-AMINOPROPYL)IMIDAZOLE-CONTAINING TRIAMIDE f-Im*PyIm: A NOVEL DIAMINO POLYAMIDE DESIGNED TO TARGET 5'-ACGCGT-3'

The work presented in this chapter is based on the published paper "*Synthesis and DNA Binding Properties of 1-(3-aminopropyl)Imidazole-Containing Triamide f-Im*PyIm: A Novel Diamino Polyamide Designed to Target 5'-ACGCGT-3'*" from *Bioorganic & Medicinal Chemistry Letters,* **2012**, accepted May 1, 2012, In press. Synthetic methods and write ups are attributed to Dr. Moses Lee of Hope College and various members of the Lee Research Group. Dr. John Hartley from the UK Drug-DNA Interactions Research Group is credited with the DNase I footprinting work and writings. My contribution to this paper was the SPR and ITC experiments, data analyses and writings, along with Dr. Yang Liu and Dr. W. David Wilson.

Synthesis and DNA binding properties of 1-(3-aminopropyl)imidazole-containing triamide f-Im*PyIm: a novel diamino polyamide designed to target 5'-ACGCGT-3'

Vijay Satam,[a] Balaji Babu,[a] Alexander Porte,[a] Mia Savagian,[a] Megan Lee,[a] Yang Liu,[b] Joseph Ramos,[b] W. David Wilson,[b] Shicai Lin,[c] Kostantinos Kiakos,[c] John Hartley,[c] and Moses Lee[a]

[a]*Department of Chemistry and the Division of Natural and Applied Sciences, Hope College 49423, USA*

[b]*Department of Chemistry, Georgia State University, Atlanta, GA, 30303, USA*

[c]*Cancer Research, UK Drug-DNA Interactions Research Group, UCL Cancer Institute, Paul O' Gorman Building, 72 Huntley Street, London WCIE 6BT, UK*

3.1 Abstract

A novel diamino/dicationic polyamide f-Im*PyIm (**5**) that contains an orthogonally positioned aminopropyl chain on an imidazole (Im*) moiety was designed to target 5'-ACGCGT-3'. The DNA binding properties of the diamino polyamide **5**, determined by CD, ΔT_M, DNase I footprinting, SPR, and ITC studies, were compared with those of its monoamino/monocationic counterpart f-ImPyIm (**1**) and its diamino/dicationic isomer f-ImPy*Im (**2**), which has the aminopropyl group attached to the central pyrrole unit (Py*). The results gave evidence for the minor groove binding and selectivity of polyamide **5** for the cognate sequence 5'-ACGCGT-3', and with strong affinity (K_{eq} = 2.3 x 10^7 M^{-1}). However, the binding affinities varied according to the order: f-ImPy*Im (**2**) > f-ImPyIm (**1**) ≥ f-Im*PyIm (**5**) confirming that the second amino group can improve affinity, but that its position within the polyamide can affect affinity.

Keywords: polyamides, diamino, dicationic, 5'ACGCGT-3', MluI cell cycle, MCB

3.2 Introduction

Pyrrole and imidazole-containing analogs of distamycin are polyamides that bind in the minor groove of DNA at specific DNA sequences in a stacked fashion.[1] They are potentially useful for targeting and modulating the expression of genes, including those associated with cancer cell growth.[2–4] Formamido(f)-Imidazole(Im)-Pyrrole(Py)-Imidazole(Im) (**1**, Figure 3.1) is such a monoamino or monocationic polyamide (PA) molecule that binds selectively as a 2:1 (polyamide:DNA) stacked dimer in the minor-groove of 5'-ACGCGT-3'. The binding affinity is 1.9×10^8 M^{-1}.[5-8] The 5'-ACGCGT-3' sequence is significant because it occurs in the core segment of the *Mlu1* cell-cycle box (MCB) transcriptional element found in the promoter of the human Dbf4 (huDbf4 or ASK, activator of S-phase kinase) gene. Dbf4 is the regulatory subunit of Cdc7 (cyclin dependent 7) kinase, and high levels of this kinase have been implicated for the development of various cancers.[9]

Despite demonstrating excellent in-vitro activity[10], the development of polyamides has been stymied by limited solubility and cellular uptake;[11] although several positive results from in-vivo experiments have been reported.[12] Our laboratories are responding to these challenges by focusing efforts on designing and testing novel polyamides that contain orthogonally positioned diamino groups. Our strategy involves functionalizing the polyamides by incorporating an additional aminoalkyl group at the *N*-position of the pyrroles or imidazoles. We hypothesized that diamino polyamides would display the following properties: enhanced binding affinity, solubility in aqueous biological media, cell penetration, nuclear localization, and at least maintains the sequence specificity. The inspiration of this work was derived from reports showing that polyamides containing a pyrrole–*N*-alkyl spermine/spermidine group targeted A/T rich sequences and bound with high affinity.[13]

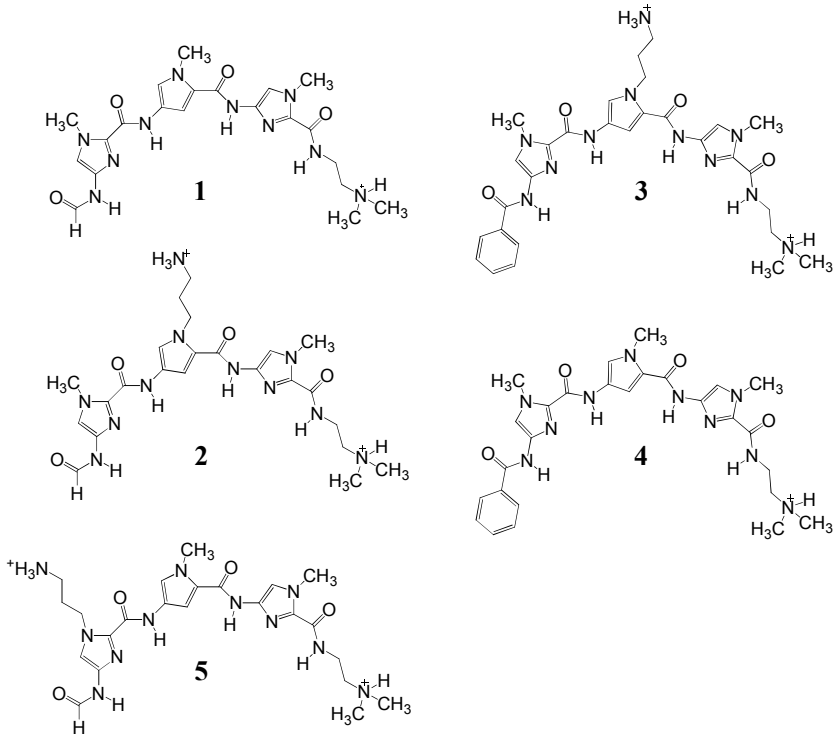
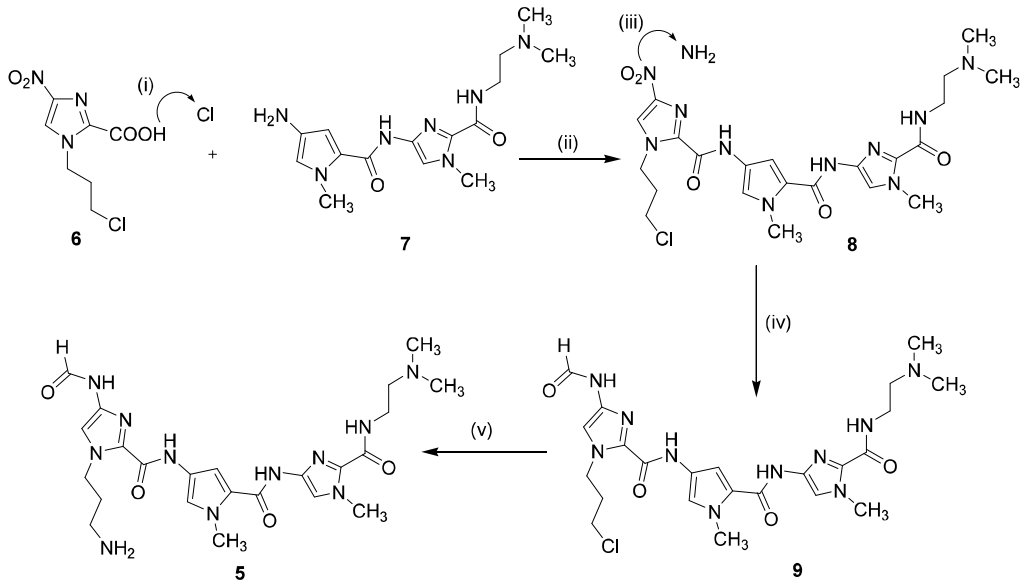


Figure 3.1 Structures of the monoamino polyamide f-ImPyIm (**1**), an orthogonally positioned diamino polyamide f-ImPy*Im (**2**), the non-formamido polyamides, diamino Ph-ImPy*Im (**3**) and monoamino Ph-ImPyIm (**4**), and a diamino polyamide f-Im*PyIm (**5**). Py* and Im* represent the *N*-(3-aminopropyl)pyrrole and 1-(3-aminopropyl)imidazole moieties, respectively.



Scheme 3.1 Reagents and conditions: (i) oxalyl chloride, dry THF, reflux, 45 min; (ii) dry triethylamine, dry DCM, rt, 18 h; (iii) H_2, 5% Pd/C, methanol, rt, 18 h; (iv) acetic formic anhydride, rt, 18 h; (iv) dry ammonia, dry methanol, 60 °C, 16 h.

3.3 Experimental

3.3.1 Synthesis

The target diamino/dicationic polyamide f-Im*PyIm **5** was synthesized according to the approach depicted in Scheme 1. 1-(3-Chloropropyl)-4-nitroimidazole-2-carboxylic acid **6**[16] and oxalyl chloride were refluxed in dry THF for 45 min and the acid chloride was coupled with freshly prepared amine **7**[5-8,17] in dry dichloromethane at room temperature. The nitro-triamide compound **8**, obtained in 36% yield, was hydrogenated over Pd-C (5%) in methanol at room temperature to yield the corresponding amine, which in turn was treated with acetic formic anhydride to afford formamido-triamide **9** in 65% yield. S_N2 displacement of the chloride in **9** with dry ammonia in methanol at 60 °C furnished the desired diamino/dicationic f-Im*PyIm polyamide **5** in 28% yield.[18]

3.3.2 DNase I footprinting studies

With f-Im*PyIm (**5**) in hand, its DNA binding properties were examined. DNase I footprinting studies were performed using a 132 bp 5'-[32P]-radiolabeled engineered DNA fragment containing the DNA sequences 5'-ACGCGT-3', 5'-ACCGGT-3', 5'-ACACGT-3' and 5'-AGCGCT-3'. The results depicted in Figure 3.2A provided direct evidence of footprinting at 5'-ACGCGT-3' emerging at 0.05 μM and becoming clearly evident at 0.10 μM. In contrast, no footprint at the three other sites was apparent until about 1 μM when some binding began to appear at 5'-AGCGCT-3'. In contrast, at 1 μM, both f-ImPyIm[5] and f-ImPy*Im[14] showed clear footprints at 5'-AGCGCT-3' and 5'ACACGT-3'. These results demonstrated a somewhat enhanced sequence specificity of f-Im*PyIm (**5**) for its cognate sequence over its parent molecule f-ImPyIm (**1**) and its isomer f-ImPy*Im (**2**). As summarized in Table 3.1, the onset of a footprint at 5'-ACGCGT-3' for f-Im*PyIm (0.05-0.10 μM) was slightly higher than that for its

parent molecule f-ImPyIm and its isomer f-ImPy*Im (both at 0.050 μM), suggesting possibly a slightly reduced binding affinity.

3.3.3 Thermal denaturation studies

Thermal denaturation experiments were performed to further examine the sequence selectivity of polyamide **5** using the following DNA sequences: 5'-ACGCGT-3', 5'-AAATTT-3' (5'-A_3T_3-3') and 5'-ATGCAT-3'. The ΔT_M values given in Table 3.1 were determined from the difference in melting temperatures of the DNA-polyamide complex and duplex DNA alone. In each case, 1.0 μM of DNA and 3.0 μM of polyamide were used. Comparable to its isomer f-ImPy*Im (**2**), the diamino polyamide f-Im*PyIm (**5**) showed remarkable interaction with its cognate sequence 5'-ACGCGT-3' as evidenced by ΔT_M values >20 °C.[14] Not surprisingly, the ΔT_M values for their non-cognate sequences were lower, particularly 5'-AAATTT-3'. These results offered evidence that a diamino polyamide that contain an aminopropyl chain attached to the 1-position of imidazole or pyrrole is able to stabilize the DNA significantly greater than its parent f-ImPyIm (**1**), which gave a ΔT_M of 7.8 °C for the cognate sequence 5'-ACGCGT-3'.[5]

3.3.4 Circular dichroism studies

CD spectroscopy was used to probe the binding of f-Im*PyIm (**5**) in the minor groove of double-stranded DNA. These experiments were conducted by titrating each polyamide with DNA solutions comprised of the DNA sequences tested in the thermal denaturation studies. As demonstrated in Figure 3.2B, the diamino polyamide **5** bound effectively to its cognate sequence 5'-ACGCGT-3'. The presence of a strong and positive induced CD band at ~330 nm in CD spectra is typical of minor grove binding by polyamides.[5-8,14,15] The distinct isodichroic point at ~310 nm suggested that **5** was bound to the DNA by a single mechanism, presumably by binding in the minor groove as a stacked dimer. This binding mode is similar to that reported for the

binding of f-ImPy*Im (**2**)[14] and the parent monoamino polyamide **1**[5] to the cognate sequence 5'-ACGCGT-3'. These results indicate that the presence of a second positively charged group in the form of an orthogonally positioned alkylammonium side chain and its position either on an imidazole or pyrrole heterocycle does not significantly affect sequence selectivity nor does it affect the conformation of the ligand-DNA complex in any significant way.

3.3.5 SPR studies

The biosensor method, SPR, has emerged as a powerful tool for measuring the binding constants of small molecules to synthetic oligonucleotides.[5-8,14,15,19] The target diamino molecule f-Im*PyIm (**5**) was tested against three oligonucleotides: the cognate sequence 5'-ACGCGT-3', and two non-cognate sequences 5'-ACATGT-3' and 5'-AAATTT-3' using this method. A representative SPR sensorgram for the binding of diamino polyamide **5** to 5'-ACGCGT-3' is depicted in Figure 3.2C. The results in Table 3.1 demonstrate that f-Im*PyIm binds more strongly to its cognate sequence than the two non-cognate sequences. As expected, it bound with the least preference for an entirely AT sequence.

Analysis of the SPR results provided evidence that polyamide **5** interacted with 5'-ACGCGT-3' with a 2:1 (polyamide:DNA) stoichiometry, i.e., as a side-by-side, stacked dimer that fit snugly in the minor groove. K_1 was calculated to be 1.8 x 10^5 M^{-1} and K_2 was 3.0 x 10^9 M^{-1}, leading to a K_{eq} of 2.3 x 10^7 M^{-1}($K_{eq} = \sqrt{K_1 \times K_2}$). In this case, the binding is a positively cooperative process. Judging from the sensorgram depicted in Figure 3.2C, it is evident that polyamide **5** interacted with 5'-ACGCGT-3' with a reasonable on-rate; however, the rate by which the polyamide dissociated from the DNA was extremely slow. Both the thermodynamic and kinetic behaviors are consistent with strong DNA binders, as was the case for polyamides **1** and **2**.[5,8,14] Further analysis of the binding of **5** with 5'-ACATGT-3', which contains two

mismatched base pairs in the core, gave values of K_1 and K_2 of 9.8 x 10^5 M^{-1} and 1.4 x 10^7 M^{-1}, respectively. The significant reduction in cooperativity compared to the binding to 5'-ACGCGT-3' is consistent with earlier observations on binding of polyamides, such as distamycin, to AT-rich sequences, such as 5'-AAATTT-3', which show negative cooperativity.[8] Presumably, the minor groove of AT-rich sequences is too narrow to comfortably accommodate two stacked polyamides.

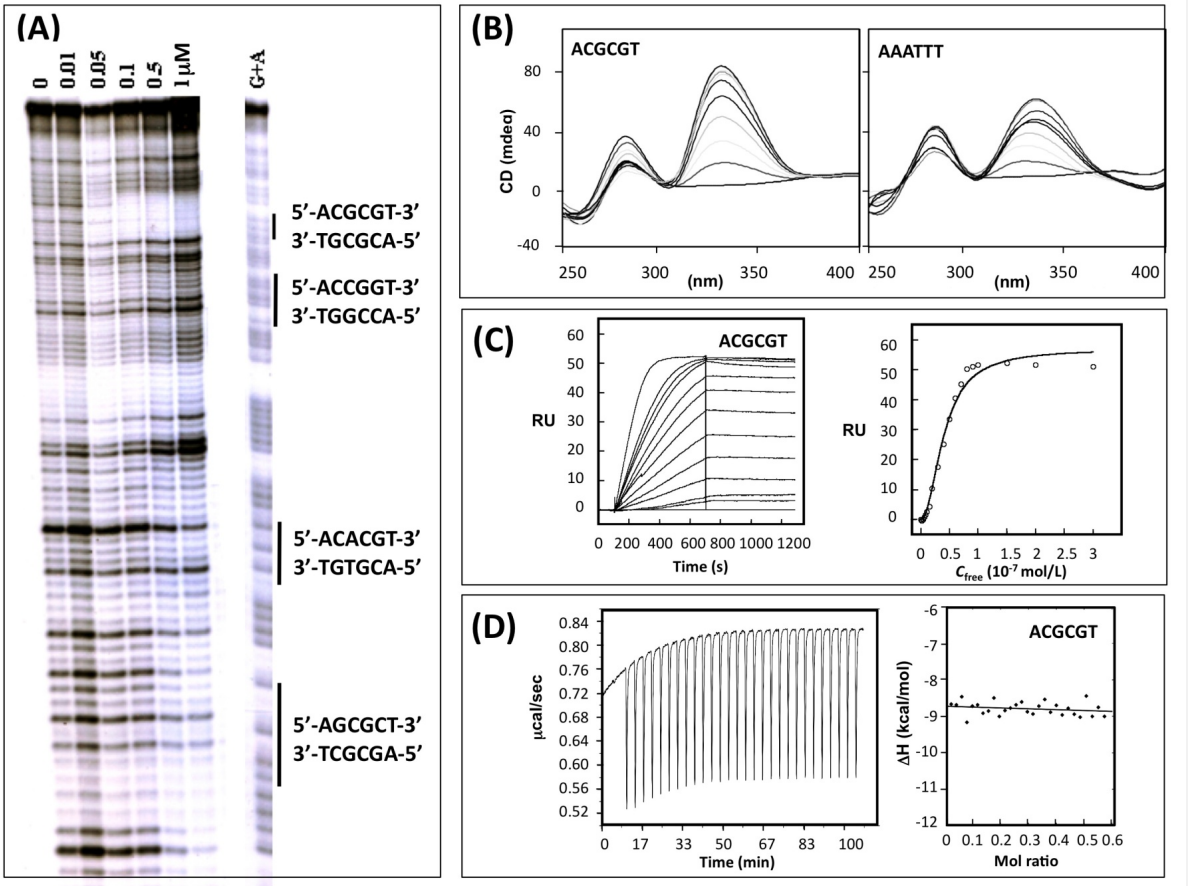


Figure 3.2 (A) DNase I footprinting of f-Im*PyIm (**5**) with 5'-ACGCGT-3', 5'-ACCGGT-3', 5'-ACACGT-3' and 5'-AGCGCT-3'. (B) Circular dichroism data for **5** with5'-ACGCGT-3' and 5'-AAATTT-3'. (C) SPR sensorgrams for **5** with cognate 5'-ACGCGT-3' and a plot of RU_{SS} *vs* concentration including best-fit lines. (D) Raw ITC data for the "model free" titration of **5** into 5'-ACGCGT-3' and the integrated heats for each injection of **5** used for the titration (black dots) and a linear fit (black line) of that data. CD experiments were carried out using 160 μL of a 9 μM DNA solution, which was titrated with 1 mol equivalents of **5,** past the point of saturation. SPR experiments were performed using concentrations of **5** from 0 - 300 nM. For the ITC studies, the titration consisted of 30-10 μL injections of 80 μM of **3** or **5** into 30 μM DNA. The titration was completed up to a mole ratio of ~0.6:1 (ligand: DNA).

3.3.6 ITC studies

To gain a deeper understanding on the binding of f-Im*PyIm (**5**) to its cognate sequence, ITC studies were conducted.[8,20] Evidence of an exothermic interaction is apparent from the data in Figure 3.2D. Integration of the heat released provided the enthalpy of interaction (ΔH) of -8.4 kcal/mol. From the binding constant determined from SPR studies, the free energy of binding (ΔG) was calculated as -10.0 kcal/mol, thereby allowing the determination of TΔS to be 1.6 kcal/mol. For comparison, the results from the binding of f-ImPy*Im (**2**) to 5'-ACGCGT-3' are also reported, and in this case the enthalpy was measured at -5.9 kcal/mol, which was much lower than that for f-Im*PyIm (**5**). The thermodynamic profile for **5** matched better with the parent molecule f-ImPyIm (**1**). The results are quite striking in demonstrating that small changes in molecular structure can significantly impact binding affinity, and at this stage it is difficult to predict whether the enthalpic or entropic term is more seriously affected.

3.4 Results and discussion

These polyamides exhibited good water solubility and they effectively inhibited the binding of transcription factors to dsDNA compared to other minor groove binders lacking an alkyl-multiamino side chain. It was suggested that upon binding, the additional cationic groups were attracted to the negatively charged phosphodiester groups of DNA.

Encouraged by the spermine/spermidine results, our group reported an orthogonally positioned diamino polyamide f-ImPy*Im (**2**), in which Py* represents the *N*-(3-aminopropyl)pyrrole moiety.[14] The results provided validation for our hypothesis. The diamino polyamide (**2**) displayed a binding constant of 2.4×10^8 M^{-1} for cognate sequence 5'-ACGCGT-3', which was four times higher than that of its monoamino/monocationic counterpart f-ImPyIm (**1**) (5.4×10^7 M^{-1}, when performed side-by-side). The sequence specificity of **2** was found to be

comparable to that of **1**. Also, incorporation of the aminopropyl group in polyamide backbone of **2** offered a significant benefit over its monocationic triamide **1** in terms of enhanced water solubility of its corresponding hydrochloride salt. With initial success in this strategy, we extended the work to explore a non-formamido polyamide, specifically diamino polyamide Ph-ImPy*Im (**3**, Figure 3.1).[15] Compared to its monoamino counterpart Ph-ImPyIm polyamide **4**, diamino polyamide **3** demonstrated an excellent level of specificity for the cognate 5'-ACGCGT-3' sequence. In addition, the binding affinity of **3** (1.5×10^7 M^{-1}) was enhanced by ~three-fold compared to its monocationic analog **4** (4.8×10^6 M^{-1}).[15] Again, the side chain of **3** was found to offer a significant benefit over its monocationic triamide **4** in terms of enhanced water solubility of the corresponding hydrochloride salt as evidenced by the ease in dissolving polyamide **3** over **4**.

It is now evident that incorporation of an orthogonally positioned aminopropyl side chain in the polyamide backbone can afford favorable DNA binding properties. Thus far, the aminoalkyl moieties have been attached to the *N*-position of pyrrole (Py*), and that is typically located in the mid-section of the polyamide structure. A question arises: how will the positioning of an aminoalkyl group at the peripheral-section of polyamides affect the DNA binding properties? In the case of the f-ImPyIm frame work, what effects will f-Im*PyIm (**5**), in which a 3-aminoalkyl group is attached to the N-terminus imidazole unit, have on DNA binding affinity, sequence selectivity, and solubility when compared to f-ImPyIm (**1**) and f-ImPy*Im (**2**). Accordingly, reported herein is the synthesis of f-Im*PyIm (**5**), which required the preparation of the key synthon, a derivatized imidazole-acid **6**.

3.5 Conclusions

Even though pyrrole and imidazole polyamides are now well known for binding strongly to DNA in a sequence specific manner, their potential realization as therapeutic agents is still in the distance. This is true particularly for the larger polyamide molecules, such as hairpins, H-Pins, and cycles.[21] Specifically, major advancements in the field that must be made are as follows: the molecules must demonstrate greater solubility in water, enhanced sequence specificity and binding affinity, and their molecular size and design must be optimized in such a way that it will enhance their penetration into cells and concentration in the nucleus. The authors' laboratories have begun to address these challenges by focusing our efforts on developing small polyamides that bind in a stacked fashion, readily enter cells, concentrate in the nucleus, and affect gene regulation. As an example, f-PyImPy, a triamide designed to target the 5'-flank (5'-TACGAT-3') of the inverted CCAAT box-2 (ICB-2) was able to inhibit the binding of a transcriptional factor NF-Y and thereby activating the silenced topoisomerase II α gene in confluent cancer cells.[22] Even though these results are encouraging, the binding affinity of f-PyImPy is low (10^5 M^{-1}) and that must be significantly improved, while simultaneously enhancing the water solubility and attraction to the DNA. The novel molecular design exemplified by diamino f-Im*PyIm (**5**), f-ImPy*Im (**2**), and Ph-ImPy*Im (**4**) offers a potential avenue for the creation of newer molecules that may possibly overcome some of the challenges in the field.

In conclusion, the synthesis of f-Im*PyIm (**5**) and knowledge of its DNA binding properties will positively advance the polyamide field in several ways. First, it was easier to dissolve the diamino compound **5** in water as a hydrochloride salt than its monoamino counterpart **1**. Second, introduction of a second positive charge in polyamides, via the 1-position

of the heterocycles, conferred positive effects on binding affinity and reducing dissociation rates of the ligand-DNA complex. Third, proper introduction of the second positive charge can improve sequence specificity, as is the case for f-Im*PyIm (**5**). Lastly, with 1-(3-chloropropyl)-4-nitroimidazole-2-carboxylic acid **6** available in addition to its pyrrole analog, we are now poised to explore more broadly the novel diamino, even multiple amino, design to optimally target specific control regions of genes in cells. Studies to examine the cellular effects of polyamides **2** and **5** in controlling the expression of the Dbf4 gene are in progress.

Table 3.1 Results from DNase I footprinting, thermal denaturation, SPR, and ITC experiments on polyamides **1**, **2** and **5**.

	Footprinting	ΔT_M (°C)[b]			SPR (M^{-1})			ITC (kcal/mol)
	(μM)[a]	ACGCGT	AAATTT	ATG CAT	ACGCGT	AAATTT	ACATGT	ACGCGT
f-ImPyIm (1)[5]	0.05	7.8	1	ND[c]	1.9 x 10^8 (5.4×10^7)[14]	5.3 x 10^4	ND[c]	ΔH = -7.6 ΔG = -10.7 TΔS = 3.1
f-ImPy*Im (2)[14]	0.05	>24	10	24	2.4 x 10^8	ND[c]	ND[c]	ΔH = -5.9 ΔG = -11.4 TΔS = 5.5
f-Im*PyIm (5)	~0.10	>20	11	20	2.3 x 10^7	2.2 x 10^6	3.7 x 10^6	ΔH = -8.4 ΔG = -10.0 TΔS = 1.6

[a]All experiments were performed using the appropriate sequence of DNA. [b]All experiments were performed in PO_4 buffer which contains 13 mM salt. [c]Not determined.

3.6 Acknowledgments

The authors thank the NSF (CHE 0809162, CHE 0550992), Cancer Research UK (C2259/A9994 to JAH) and the Georgia Research Alliance for their generous support.

3.7 References and notes

1. Mitra, S. N.; Wahl, M. C.; Sundaralingam, M. *Acta Crystallogr. Sect. D* **1999**, *55*, 602; (b) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci.* **1989**, *86*, 5723; (c) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc*. **1990**, *112*, 1393.
2. (a) Chou, C. J.; Farkas, M. E.; Tsai, S. M.; Alvarez, D.; Dervan, P. B.; Gottesfeld, J. M. *Mol. Cancer Ther.* **2008**, *7*, 769; (b) Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Nat. Struct. Biol.* **1996**, *3*, 321; (c) Chenoweth, D. M.; Dervan, P. B. *Proc. Natl. Acad. Sci.* **2009,** *106*, 13175.
3. (a) Hochhauser, D.; Kotecha, M.; O'Hare, C.; Morris, P. J.; Hartley, J. M.; Taherbhai, Z.; Harris, D.; Forni, C.; Mantovani, R.; Lee, M.; Hartley, J. A. *Mol. Cancer Ther.* **2007**, *6*, 346; (b) Dwyer, T. J.; Geierstanger, B. H.; Bathini, Y.; Lown, J. W.; Wemmer, D. E. *J. Am. Chem. Soc.* **1992,** *114*, 5911.
4. (a) Shinohara, K.; Bando, T.; Sugiyama, H. *Anticancer Drugs* **2010**, *21*, 228; (b) Minoshima, M.; Bando, T.; Shinohara, K.; Sugiyama, H. *Nucleic Acids Symp. Ser.* **2009**, *23*, 69; (c) Borman, S. *Chem. Eng. News* **2010**, *88*, 50.
5. Buchmueller, K. L.; Bailey, S. L.; Matthews, D. A.; Taherbhai, Z. T.; Register, J. K.; Davis, Z. S.; Bruce, C. D.; O'Hare, C.; Hartley, J. A.; Lee, M. *Biochemistry* **2006**, *45*, 13551.
6. (a) Lacy, E. R.; Le, N. M.; Price, C. A.; Lee, M.; Wilson, D. W. *J. Am. Chem. Soc.* **2002**, *124*, 2153; (b) Westrate, L.; Mackay, H.; Brown, T.; Nguyen, B.; Kluza, J.; Wilson, D. W.; Lee, M.; Hartley, J. A. *Biochemistry* **2009**, *48*, 5679.
7. Buchmueller, K. L.; Staples, A. M.; Uthe, P. B.; Howard, C. M.; Pacheco, K. A.; Cox, K. K.; Henry, J. A.; Bailey, S. L.; Horick, S. M.; Nguyen, B.; Wilson, D. W.; Lee, M. *Nucleic Acids Res.* **2005**, *33*, 912.
8. Buchmueller, K. L.; Staples, A. M.; Howard, C. M.; Horick, S. M.; Uthe, P. B.; Le, N. M.; Cox, K. K.; Nguyen, B.; Pacheco, K. A.; Wilson, D. W.; Lee, M. *J. Am. Chem. Soc.* **2005**, *127*, 742.
9. (a) Verma, R.; Patapoutian, A.; Gordon, C. B.; Campbell, J. L. *Proc. Natl. Acad. Sci.* **1991**, *88*, 7155; (b) Wu, X.; Lee, H. *Oncogene* **2002**, *21*, 7786; (c) Yamada, M.; Sato, N.; Taniyama, C.; Ohtani, K.; Drong, R. F.; Weiland, K. L.; Slightom, J. L.; Sclafani, R. A.; Hollingsworth, R. E. *Gene* **1998**, *28*, 133.
10. (a) Jacobs, C. S.; Dervan, P. B. *J. Med. Chem*. **2009**, 52, 7380; (b) Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *ACS Chem. Biol*. **2007**, 2, 561; (c) Nickols, N. G.; Dervan, P.

B. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, 104, 10418; (d) Burnett, R.; Melander, C.; Puckett, J. W.; Son, L. S.; Wells, R. D.; Dervan, P. B.; Gottesfeld, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 11497; (e) Puckett, J. W.; Muzikar, K. A.; Tietjen, J.; Warren, C. L.; Ansari, A. Z.; Dervan, P. B. *J. Am. Chem. Soc*. **2007**, 129, 12310.

11. (a) Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007**, *35*, 363; (b) Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res*. **2004**, *32*, 2802; (c) Hsu, C. F.; Dervan, P. B. *Bioorg. Med. Chem. Lett*. **2008**, *18*, 5851; (d) Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 12063.

12. (a) Chou, C. J.; Farkas, M. E.; Tsai, S. M.; Alvarez, D.; Dervan, P. B.; Gottesfeld, J. M. *Mol Cancer Ther.* **2008**, *7*, 769; (b) Harki, D. A.; Satyamurthy, N.; Stout, D. B.; Phelps, M. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 13039; (c) Fukasawa, A.; Aoyama, T.; Nagashima, T.; Fukuda, N.; Ueno, T.; Sugiyama, H.; Nagase, H.; Matsumoto, Y. *Biopharm. Drug Dispos*. **2009**, *30*, 81; (d) Matsuda, H.; Fukuda, N.; Ueno, T.; Tahira, Y.; Ayame, H.; Zhang, W.; Bando, T.; Sugiyama, H.; Saito, S.; Matsumoto, K.; Mugishima, H.; Serie, K. *J. Am. Soc. Nephrol.* **2006**, *17*, 422.

13. (a) Satz, A. L.; Bruice, T. C. *Acc. Chem. Res.* **2002**, *35*, 86; (b) White, C. M.; Satz, A. L.; Gawron, L. S.; Bruice, T. C.; Beerman, T. A. *Biochim. Biophys. Acta*. **2002**, *1574*, 100; (c) Satz, A. L.; Bruice, T. C. *Bioorg. Med. Chem*. **2002**, *10*, 241; (d) White, C. M.; Satz, A. L.; Bruice, T. C.; Beerman, T. A. *Proc. Natl. Acad. Sci. U. S. A*. **2001**, 98, 10590.

14. Babu, B.; Liu, Y.; Plaunt, A.; Riddering, C.; Ogilvie, R.; Westrate, L.; Davis, R.; Ferguson, A.; Mackay, H.; Rice, T.; Ramos, J. P.; Chavda, S.; Wilson, D.; Lin, S.; Kiakos, K.; Hartley, J. A.; Lee, M. *Biochem. Biophys. Res. Commun.* **2011**, *404*, 848.

15. Satam, V.; Babu, B.; Chavda, S.; Savagian, M.; Sjoholm, R.; Tzou, S.; Ramos, J.; Liu, Y.; Kiakos, K.; Lin, S.; Wilson, D. W.; Hartley, J. A.; Lee, M. *Bioorg. Med. Chem.* **2012**, *20*, 693.

16. The synthesis of 1-(3-chloropropyl)-4-nitroimidazole-2-carboxylic acid **6** will be published elsewhere.

17. Chavda, S.; Liu, Y.; Babu, B.; Davis, R.; Sielaff, A.; Ruprich, J.; Westrate, L.; Tronrud, C.; Ferguson, A.; Franks, A.; Tzou, S.; Adkins, C.; Rice, T.; Mackay, H.; Kluza, J.; Tahir, S. A.; Lin, S.; Kiakos, K.; Bruce, C. D.; Wilson, D. W.; Hartley, J. A.; Lee, M. *Biochemistry* **2011**, *50*, 3127.

18. Diamino polyamide **5** was isolated as a white solid (15 mg, 28%), mp 83 °C, R_f 0.14 (1:1 v/v, $CHCl_3$:CH_3OH); IR: (υ) 2962, 1528, 1516, 1464, 1445, 1260, 1089, 1018, 971, 918, 882, 870, 846, 796, 748, 725, 697 cm^{-1}; ^1H NMR (CD_3OD) 8.26 (s, 1H); 7.51 (s, 1H); 7.42 (s, 1H); 7.35 (s, 1H); 7.00 (s, 1H); 4.58 (t, *J*=6.0 Hz, 2H); 4.01 (s, 3H); 3.94 (s, 3H); 3.48 (t, *J*=6.0 Hz, 2H); 2.68 (t, *J*=6.0 Hz, 2H); 2.55 (t, *J*=4.0 Hz, 2H); 2.31 (s, 6H); 2.03 (q, *J*=6.0 Hz, 2H); LRMS (TOF-ES$^+$) *m/z* (rel. intensity) 528 ($M+H^+$, 25%), 265 [$(M+2H)^{2+}$, 100%]. ESI-TOF-HRMS $[M+H]^+$ calcd for *m/z* $C_{23}H_{34}N_{11}O_4^+$ 528.2795, found 528.2809.

19. (a) Wilson, W. D. *Science* **2002***, 295*, 2103; (b) Liu, Y.; Wilson, W. D. *Methods Mol. Biol.* **2010**, *613*, 1; (c) SPR measurements were performed with a four-channel BIAcore T200 optical biosensor system (Biacore, GE Healthcare Inc.). 5'-Biotin-labeled DNA hairpin duplex samples were immobilized onto streptavidin-coated sensor chips (Biacore SA) as previously described.[5-8,14,15] All SPR experiments were performed at 25 °C and used a 0.01 M

CCA, 100 mM NaCl, pH 6.2 buffer. The amount of DNA immobilized was approximately 400 response units (RU). This was achieved by continuously injecting ~20 μL of an approximately 50 nM DNA solution at a rate of 2 μL/min onto the sensor chip surface until a relative response of 400 units was reached. Binding data were obtained by injecting known concentrations and were analyzed with one or two site binding models as previously described: $r = RU/RU_{max} = (K_1 \times (C_{free} + 2 \times K_1 \times K_2 \times C^2_{free})) / (1 + K_1 \times C_{free} + K_1 \times K_2 \times C^2_{free})$ where, r represents the moles of bound compound per mole of DNA hairpin duplex, RU is the steady-state response, K_1 and K_2 are macroscopic binding constants, and C_{free} is the free compound concentration in equilibrium with the complex. For curves that did not reach a steady-state, the RU value was determined by extrapolation.

20. (a) Indyk, L.; Fisher, H. F. *Methods Enzymol*. **1998**, *295*, 350; (b) Lacy, E. R.; Nguyen, B.; Le, M.; Cox, K. K.; O'Hare, C.; Hartley, J. A.; Lee, M.; Wilson, W. D. *Nucleic Acids Res.* **2004**, *32*, 2000; (c) ITC analysis was performed using a VP ITC microcalorimeter (MicroCal) and the excess DNA or model free method. The instrument was equilibrated at the noted temperature and after an initial delay of 300 s, compound **5** was titrated into DNA. Origin 7.0 was used in data analysis and the area under the curve integrated as a function of time. A linear fit was then employed to determine the binding enthalpy, and this was subtracted from the reaction integrations to normalize for nonspecific heat components. ΔG was calculated from $\Delta G = -RT \ln K_{eq}$, where, R is 1.987 cal mol^{-1} K^{-1} and T is measured in K.

21. a) Dervan, P. B.; Doss, R. M.; Marques, M. A. *Curr. Med. Chem.: Anti-Cancer Agents* **2005**, *5*, 373; (b) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struc. Biol*. **2003**, *13*, 284; (c) Dervan, P. B. *Bioorg. Med. Chem*. **2001**, *9*, 2215.

22. Le, N. M.; Sielaff, A. M.; Cooper, A. J.; Mackay, H.; Brown, T.; Kotecha, M.; O'Hare, C.; Hochhauser, D.; Lee, M.; Hartley, J. A. *Bioorg. Med. Chem. Lett*. **2006**, *16*, 6161.

CHAPTER 4:

NOVEL DIAMINO IMIDAZOLE AND PYRROLE-CONTAINING POLYAMIDES: SYNTHESIS AND DNA BINDING STUDIES OF MONO- AND DIAMINO-PHENYL-ImPy*Im POLYAMIDES DESIGNED TO TARGET 5'-ACGCGT-3'

The work presented in this chapter is based on the published paper "*Novel diamino imidazole and pyrrole-containing polyamides: Synthesis and DNA binding studies of mono- and diamino-phenyl-ImPy*Im polyamides designed to target 5'-ACGCGT-3'"* from *Bioorganic and Medicinal Chemistry*, **2012**, 20, 693-701. Synthetic methods and write ups are attributed to Dr. Moses Lee of Hope College and various members of the Lee Research Group. Dr. John Hartley from the UK Drug-DNA Interactions Research Group is credited with the DNase I footprinting work and writings. My contribution to this paper was the SPR experiments, data analyses and writings, along with Dr. Yang Liu and Dr. W. David Wilson.

Novel diamino imidazole and pyrrole-containing polyamides: Synthesis and DNA binding studies of mono- and diamino-phenyl-ImPy*Im polyamides designed to target 5'-ACGCGT-3'

Vijay Satam,[a] Balaji Babu,[a] Sameer Chavda,[a] Mia Savagian,[a] Robert Sjoholm,[a] Samuel Tzou,[a] Joseph Ramos,[b] Yang Liu,[b] Konstantinos Kiakos,[c] Shicai Lin,[c] W. David Wilson,[b] John Hartley,[c] Moses Lee[a*]

[a] *Division of Natural & Applied Sciences and Department of Chemistry, Hope College, 35 East, 12th Street, Holland, MI 49423, USA*

[b] *Department of Chemistry, Georgia State University, Atlanta, Georgia 30302, USA*

[c] *Cancer Research UK Drug-DNA Interactions Research Group, UCL Cancer Institute, Paul O'Gorman Building, 72 Huntley Street, London WC1E 6BT, UK*

4.1 Abstract

Pyrrole- and imidazole-containing polyamides are widely investigated as DNA sequence selective binding agents that have potential use as gene control agents. The key challenges that must be overcome to realize this goal is the development of polyamides with low molar mass so the molecules can readily diffuse into cells and concentrate in the nucleus. In addition, the molecules must have appreciable water solubility, bind DNA sequence specifically, and with high affinity. It is on this basis that the orthogonally positioned diamino/dicationic polyamide Ph-ImPy*Im **5** was designed to target the sequence 5'-ACGCGT-3'. Py* denotes the pyrrole unit that contains a *N*-substituted aminopropyl pendant group. The DNA binding properties of diamino polyamide **5** were determined using a number of techniques including CD, ΔT_M, DNase I footprinting, SPR and ITC studies. The effects of the second amino moiety in Py* on DNA binding affinity over its monoamino counterpart Ph-ImPyIm **3** were assessed by conducting DNA binding studies of **3** in parallel with **5**. The results confirmed the minor groove binding and selectivity of both polyamides for the cognate sequence 5'-ACGCGT-3'. The diamino/dicationic polyamide **5** showed enhanced binding affinity and higher solubility in aqueous media over its monoamino/monocationic counterpart Ph-ImPyIm **3**. The binding constant of **5**, determined from SPR studies, was found to be 1.5×10^7 M^{-1}, which is ~3 times higher than that for its monoamino analog **3** (4.8×10^6 M^{-1}). The affinity of **5** is now approaching that of the parent compound f-ImPyIm **1** and its diamino equivalent **4**. The advantages of the design of diamino polyamide **5** over **1** and **4** are its sequence specificity and the ease of synthesis compared to the *N*-terminus pyrrole analog **2**.

Keywords: Polyamides, diamino, dicationic, gene control, ACGCGT, MCB, *Mlu1*, cell-cycle box

4.2 Introduction

Pyrrole and imidazole-containing analogs of distamycin are polyamides that bind in the minor-groove of DNA in a stacked fashion at specific DNA sequences.[1] They are potentially useful for targeting and modulating the expression of genes, including those associated with cancer cell growth.[2–4] Formamido(f)-Imidazole(Im)-Pyrrole(Py)-Imidazole(Im) (**1**, Figure 4.1) is one such polyamide (PA) molecule that selectively binds 5'-ACGCGT-3' sequence in the minor-groove of DNA in a 2:1 fashion (PA:DNA), with a binding affinity of 1.9×10^8 M^{-1}.[5] The 5'-ACGCGT-3' sequence is significant because it occurs in the core sequence of the *Mlu1* cell-cycle box (MCB) transcriptional element found in the promoter of the human DBf4 (huDbf4 or ASK, activator of S-phase kinase) gene. Dbf4 is the regulatory subunit of Cdc7 (cyclin dependent 7) kinase, and high levels of this kinase have been implicated for the development of various cancers.[6] The formamido group at the *N*-terminus of such polyamides has been shown to influence the binding of these compounds by forcing them to stack in a 'staggered', rather than an 'overlapped' motif.[7]

Polyamides in which the formamido group is replaced with a pyrrole or aryl groups at the *N*-terminus are known to bind in the 'overlapped' fashion; however, with diminished binding affinity but increased sequence selectivity.[8] The non-formamido polyamide (**2**, Figure 4.1), which contains a pyrrole moiety at the *N*-terminus was synthesized while keeping the C-terminus dimethylamino moiety and heterocyclic core (-ImPyIm-) constant.[8] It was found to retain its preference for binding to 5'-ACGCGT-3', but with enhanced sequence specificity albeit with a 25-fold lower binding affinity (7.1×10^6 M^{-1}) compared to f-ImPyIm (**1**) (1.9×10^8 M^{-1}).[8] This is not surprising since the formamido group is known to confer enhanced binding affinity.[7a,9] With regard to enhanced sequence selectivity, whilst f-ImPyIm (**1**) gave a binding affinity of 2.2×10^5

and 5.3×10^4 M^{-1} for 5'-ACCGGT-3' and 5'-AAATTT-3', respectively,[5] PyImPyIm (**2**) did not produce binding to these two sequences at the same concentrations.[8] It is apparent from these results that tetraamide-based structures, such as **2,** are excellent templates for the design of novel DNA sequence specific agents for biological activity, yet having acceptable binding affinity. However, in our hands the synthesis of PyImPyIm (**2**), with a pyrrole moiety at the *N*-terminus, was cumbersome owing to susceptibility of *N*-methylpyrrole-2-carbonyl chloride to undergo polymerization during the reaction.[10] Our approach to overcome this limitation is to synthesize a novel polyamide Ph-ImPyIm (**3**, Figure 4.1) by incorporating a phenyl moiety, in place of pyrrole, in tetraamide **2**. We envisaged that incorporation of a phenyl group, instead of pyrrole, at the *N*-terminus would simplify the process due to the higher stability and ready availability of inexpensive benzoyl chloride. It was also anticipated that the phenyl component of such a polyamide would behave similarly to a non-formamido, *N*-terminal pyrrole unit (Py), thus binding to A/T base pairs. In addition, benzamides have been reported to function in a similar manner to pyrrole-2-carboxamides in recognizing AT-sequences of DNA.[11,12] To confirm this hypothesis, a hitherto unreported novel polyamide, Ph-ImPyIm (**3**) was synthesized and evaluated for its DNA binding properties.

As part of a systematic study within the authors' laboratory our goal is to develop MCB-targeted polyamides which can be synthesized readily yet possess excellent sequence specificity, stronger binding affinity, high solubility in biological media and enhanced cell penetration and nuclear localization properties. The strategy undertaken in our laboratory to increase the solubility of polyamides is inclusion of an additional amine/ammonium group in the polyamide structure without compromising sequence selectivity. In this strategy, the authors are aware of earlier reports indicating that imidazole-containing polyamides with multiple cationic groups

have affinity for A/T rich sequences due to attraction to the negative molecular electrostatic potential in the minor groove of A/T rich sequences.[13,14] Hence, caution must be exercised to develop multiamino polyamides. In this regard, Bruice and Satz have reported that polyamides containing a pyrrole–N1-alkyl spermine/spermidine group target A/T rich sequences and bind with high affinity.[15] These polyamides effectively inhibit the binding of transcription factors to dsDNA compared to other minor groove binders lacking an alkyl-multiamino side chain. It was suggested that upon binding, the additional cationic groups were attracted to the negatively charged phosphodiester groups of DNA. The presence of a second amino group (cationic at physiological pH) into the polyamide design should allow greater penetration in cells due to an increase in solubility of the compounds in biological media. Encouraged by these precedents, an orthogonally positioned diamino polyamide f-ImPy*Im (**4**) was synthesized in the authors' laboratory.[16] Investigation of DNA binding properties revealed that **4** had a binding constant that is 4 times higher than that of its monoamino/monocationic counterpart f-ImPyIm (**1**) (5.4×10^7 M^{-1}). The sequence specificity of **4** was found to be comparable to that of **1**. Incorporation of the propylamino group in polyamide backbone of **4** also offered a significant benefit over its monocationic triamide **1** in terms of enhanced water solubility of its corresponding HCl salt. Hence, to exploit the benefits of having a second positively charged group in the form of an orthogonally positioned alkylammonium side chain, the diamino or potentially dicationic Ph-ImPy*Im (**5**, Figure 4.1) was designed and synthesized, in which Py* represents the *N*-(2-aminopropyl)pyrrole moiety. A comparative study of DNA recognition properties of diamino containing polyamide **5** with that of monoamino containing polyamide **3** is presented in Figure 4.2.

4.3 Results and discussion

4.3.1 Synthesis

The monoamino Ph-ImPyIm **3** and the orthogonally positioned diamino/dicationic polyamide **5** were synthesized according to the approach depicted in Schemes 4.1 and 4.2, respectively. Hydrogenation of ethyl 1-methyl-4-nitroimidazol-2-carboxylate[17] **6** with 5% palladium on charcoal in ethanol followed by coupling of the resulting amine with benzoyl chloride in dry DCM in the presence of dry triethylamine gave ethyl 4-benzamido-1-methylimidazol-2-carboxylate **7** in 96% yield. The ester **7** was hydrolyzed with sodium hydroxide in a methanol: water (1:1) mixture at reflux to obtain the corresponding carboxylic acid 4-benzamido-1-methylimidazole-2-carboxylic acid **8** which was coupled to amino-pyrrole-imidazole diamide **9**[18] using PyBOP and diisopropylethylamine in DMF at room temperature. The reaction afforded the monoamino Ph-ImPyIm (**3**) in 38% yield. Similar coupling of the acid **8** with amino containing diamide **10**[16] yielded chloro compound **11**, which upon treatment with dry ammonia in methanol at 80 °C furnished the desired diamino/dicationic polyamide **5** in 62% yield.

4.3.2 DNase I footprinting studies

The sequence specificity of the monoamino polyamide **3** and its diamino counterpart **5** was investigated by DNase I footprinting studies[19] using a 125 bp 5'-[^{32}P]-radiolabeled engineered DNA fragment containing the cognate sequence 5'-ACGCGT-3', and the non-cognate 5'-ACCGGT-3', 5'-ACACGT-3' and 5'-AGCGCT-3'. The autoradiogram depicted in Figure 4.3A shows that the experiment with Ph-ImPyIm (**3**) produced a footprint at the 5'-

ACGCGT-3' site at 20 μM, and no other footprints were apparent even at 100 μM. The results for diamino polyamide **5** demonstrated superior binding affinity compared to its monoamino counterpart **3**. The footprint for 5'-ACGCGT-3' for diamino polyamide **5** emerged at 0.5 μM (Fig. 3B), which is about 10-fold lower than its formamido parent compound f-ImPyIm (**1**).[5] This is an exciting discovery because having the second amino group made the molecule easier to dissolve in aqueous solution; it has an increased binding affinity, and its binding constant for 5'-ACGCGT-3' compared to Ph-ImPyIm (**3**) is increased by a factor of about 3. In comparison to f-ImPyIm (**1**), the binding constant of diamino polyamide Ph-ImPy*Im (**5**) was only 10-fold lower. Overall, the gain in binding affinity of compound **5** over **3** was achieved without compromising on sequence selectivity.

4.3.3 Thermal denaturation studies

Thermal denaturation experiments were performed to further examine the binding of polyamides **3** and **5** using following DNA sequences; 5'-ACGCGT-3', 5'-AAATTT-3' (5'-A_3T_3-3'), and 5'-ACACTT-3'. These thermal denaturation experiments provided an indication on the preferred binding of the monoamino polyamide **3** along with its diamino/dicationic counterpart **5** to the DNA by measuring their ability to stabilize the DNA duplex upon heating. ΔT_M values were determined from the differences in melting temperature of the DNA-polyamide complexes and duplex DNA alone and are shown in Table 4.1. In each case, 1.0 μM of DNA and 3.0 μM of polyamide were used. Comparable to its parent polyamide f-ImPyIm (**1**),[5] the monoamino polyamide **3** showed good binding interaction with the cognate sequence 5'-ACGCGT-3' by giving a ΔT_M of 9 °C. It did not stabilize the other DNA sequences used in this study as demonstrated by the ΔT_M values expressed in Table 4.1. These results indicate that the monoamino polyamide **3** possesses selectivity for its cognate DNA sequence 5'-ACGCGT-3'.

For the diamino Ph-ImPy*Im, the ΔT_M value for binding to its cognate sequence 5'-ACGCGT-3' is remarkable at >25 °C. It is able to stabilize the DNA significantly greater than either its monocationic counterpart **3** or its parent f-ImPyIm (**1**)[5], both of which gave a ΔT_M value of 9 °C.

4.3.4 Circular dichroism studies

CD spectroscopy was used to probe the binding of polyamides **3** and **5** in the minor groove of double-stranded DNA. These experiments were conducted by titrating each polyamide with DNA solutions comprised of the DNA sequences tested in the thermal denaturation studies. In all cases, a fixed DNA concentration of 9 μM was used as were the ratios of polyamides (1, 2, 3, 4, 5, 6, and 8 molar equivalents) titrated into the solution. Both polyamides **3** and **5** bind effectively to their cognate sequence 5'-ACGCGT-3'as can be seen from Figure 4.4. The presence of a strong and positive induced CD band at ~330 nm in CD spectra shown in Fig. 4 suggests minor grove binding for polyamides **3** and **5**. The distinct isodichroic point at ~310 nm provides evidence that both **3** and **5** bind to the DNA sequence by a single mechanism, presumably by interacting in the minor groove as a stacked dimer. These results also indicate that an increase of one positive charge does not significantly affect sequence selectivity. The results also show that polyamide **5** could bind to its non-cognate sequence at the concentrations used in CD experiments and exhibit some tolerance for flanking base pair on either side of the central GC core. These results are corroborated by the aforementioned ΔT_M values for **3** and **5** (Table 4.1).

4.3.5 Biosensor-surface plasmon resonance studies

To obtain a more accurate measure of binding affinity and sequence selectivity plus to probe the stoichiometry of binding, polyamides **3** and **5** were subjected to experiments using the

surface plasmon resonance (SPR)-biosensor method.[20] The experiments were conducted using cognate DNA sequence 5'-ACGCGT-3' and non-cognate sequences 5'-AAGCTA-3' and 5'-AAATTT-3'. The sensorgrams recorded from the SPR experiments for **3** and **5** binding to 5'-ACGCGT-3' are given in Figure 4.5. According the sensorgram pattern displayed in Figure 4.5B , it is clearly evident that the monoamino polyamide **3** binds strongly to its cognate sequence 5'-ACGCGT-3' as a cooperative dimer (n~2) with a binding constant (K_{eq}) 4.8×10^6 M^{-1}. The sensorgram shown in Figure 5.5A displays a similar slow observed kinetics pattern suggesting that the diamino polyamide **5** also binds strongly to the cognate sequence 5'-ACGCGT-3' as a cooperative dimer (n~2.2). The binding constant (K_{eq}) of **5** to its cognate sequence was found to be 1.5×10^7, which is almost 3-fold higher than that of **3**. This result is consistent with the results recorded for the DNase I footprinting experiments. Both polyamides **3** and **5** also bind to their non-cognate DNA sequence 5'-AAGCTA-3', but with a lower affinity (Table 4.2). The monoamino Ph-ImPyIm (**3**) binds to 5'-AAGCTA-3' as a dimer in a non-cooperative manner with binding constant (K_{eq}) 2.5×10^5 M^{-1}, while its diamino counterpart **5** binds to same DNA sequence also as a dimer but in a slightly cooperative manner with higher binding constant (K_{eq}) 2.1×10^6 M^{-1}. Binding to 5'-AAGCTA-3' is not surprising since the stacked dimer of both polyamides **3** and **5** contain the "ImPy" central pairing, which the authors have reported to give the most favorable binding to the "GC" core sequence.[5] However, none of the polyamides showed any binding affinity to their non-cognate DNA sequence 5'-AAATTT-3' under these conditions. This is not surprising since all previously reported f-ImPyIm analogs[5,16] did not show any binding to the 5'-AAATTT-3' (data not shown), indicating GC-sequence preference by Im/Py pairs. Furthermore, the enhanced binding of **5** over **3** is evidently a result of a slightly slower or "flatter" off-rate as revealed by the SPR sensorgrams for dicationic polyamide **5** than

those for the monoamino analog **3**. To further elaborate on the sequence selectivity of non-formamido polyamides, such as the phenyl-containing polyamides **3** and **5**, the SPR data given in Table 4.2 were compared to those previously reported for f-ImPyIm (**1**). It is evident that whilst polyamides **3** and **5** did not show any binding to 5'-AAATTT-3', f-ImPyIm (**1**) was reported to bind and it gave a binding constant of 5.3 x 10^4 M^{-1}, albeit it was almost 4-orders of magnitude lower than the binding of compound **1** to its cognate 5'-ACGCGT-3'. This observation provided evidence for enhanced sequence selectivity of the phenyl-containing polyamides **3** and **5** compared to their formamido counterpart **1**.

4.3.6 Isothermal titration calorimetry studies

ITC experiments were performed at 25 °C to probe the thermodynamics of DNA binding interactions of **3** and **5**. The monoamino polyamide **3** again demonstrated binding to the cognate sequence as evidenced by the exothermic enthalpy of binding of **3** to 5'-ACGCGT- 3' (Figure 4.6). Because of the very strong binding of these polyamides to the cognate DNA, a constant enthalpy was obtained below a mole ratio of 0.6 (Figure 4.6) where essentially all of the added compound is bound to DNA. A linear fit and extrapolation back to 0.0 mole ratio gave a ΔH of ~ -2.2 ± 0.14 kcal/mol. The similar binding interaction with 5'-ACGCGT-3'was observed for diamino polyamide **5** as evident from the ITC data shown in Fig. 6. A linear fit and extrapolation of heats back to 0.0 mole ratio yielded a ΔH of ~ -2.7 ± 0.2 kcal/mol (Fig. 6). With these ΔH values and the ΔG for binding from SPR ($\Delta G = -RT \ln K_{eq}$ where, R is 1.987 cal mol^{-1} K^{-1} and T is measured in K) the TΔS values were calculated (ΔG = ΔH - TΔS). The results for 5'-ACGCGT-3' reveal that the enhanced binding affinity, hence the free energy of binding, in the dicationic polyamide **5** over monocation **3** is driven more significantly by an enthalpy gain (ΔΔH ~ 0.5 kcal/mol). In contrast, the gain in entropy was smaller (ΔTΔS ~ 0.2 kcal/mol). The results

suggest that introduction of a second amino moiety provides favorable thermodynamic interactions between the ammonium function and DNA, presumably via electrostatic or hydrogen bond interactions with the phosphate group in the DNA backbone.

An even more drastic difference is noted by comparing the thermodynamic data of polyamides **3** and **5** to their parent molecule f-ImPyIm (**1**).[5] The results given in Table 4.2 show that unlike the phenyl-containing polyamide whose binding was driven significantly by entropic means, the binding of f-ImPyIm (**1**) was driven primarily by enthalpy. This finding is intriguing and useful because it demonstrated that even a small change from *N*-formamido to an *N*-benzamide could produce dramatic changes in the thermodynamics of interactions. The results indicate that in our next step of molecular design to further increase the binding affinity of diamino compound **5**, the strategy must include elements that increase enthalpy. Such elements would include incorporating groups that would increase favorable intermolecular interactions.

4.4 Conclusion

The results from DNA binding studies show that having a second positively charged group in the form of an orthogonally positioned alkylammonium side chain on the polyamide backbone does not affect DNA sequence selectivity. In fact, the diamino polyamide Ph-ImPy*Im (**5**) demonstrated an excellent level of specificity for the cognate 5'-ACGCGT-3' sequence compared to its monoamino counterpart **3**. In addition, the binding affinity of **5** is enhanced by ~3-fold compared to its monocationic analog **3**; thus, effectively raising the binding affinity of **5** to 5'-ACGCGT-3' to a level that is only 10-fold lower than its parent molecule f-ImPyIm (**1**). The enhanced binding affinity of **5** over **3** is possibly due to favorable electrostatic interaction between the negatively charged DNA phosphodiester groups and the ammonium

species. The side chain of **5** also offers a significant benefit over its monocationic triamide **3** in terms of enhanced water solubility of the corresponding HCl salts as evidenced by the ease in dissolving polyamide **5** over **3**. Given its small size, it is our hypothesis that diamino polyamides such as **5** should readily be taken up by cells. Further studies to probe the ability of polyamides **3** and **5** to affect gene function are underway and results will be reported in due course.

4.5 Experimental

Solvents and organic reagents were purchased from Aldrich or Fisher, and in most cases were used without further purification. DCM (P_2O_5), and DMF (BaO) were distilled prior to use. Melting points (mp) were performed using a Mel-temp instrument and are uncorrected. Infrared (IR) spectra were recorded using a Perkin Elmer Paragon 500 FT-IR instrument as films on KBr disks. ^{1}H NMR spectra were obtained using a Varian Unity Inova 400 MHz instrument unless otherwise stated. Chemical shifts (δ) are reported at 20 °C in parts per million (ppm) downfield from internal tetramethylsilane (Me_4Si). High-resolution mass spectra (HRMS) and low-resolution mass spectra (LRMS) were provided by the Mass Spectrometry Laboratory, University of South Carolina, Columbia. Reaction progress was assessed by thin-layer chromatography (TLC) using Merck silica gel (60 F_{254}) on aluminum plates unless otherwise stated. Visualization was achieved with UV light at 254 nm and/or 366 nm, I_2 vapor staining and ninhydrin spray.

4.6 Synthesis

Ethyl 4-benzamido-1-methylimidazole-2-carboxylate (7). Ethyl 1-methyl-4-nitroimidazole-2-carboxylate[17] (**6**, 613 mg, 3.1 mmol) was hydrogenated over 5% palladium on charcoal in ethanol. The resulting amino compound was dried thoroughly under vacuum to remove traces of

ethanol. It was then dissolved in dichloromethane (20.0 mL) and the solution was cooled to 0-5 °C. To the cold solution was added triethylamine (0.5 mL, 3.7 mmol) followed by benzoyl chloride (0.4 mL, 3.4 mmol). The reaction mixture was allowed to attain room temperature and then stirred for 12 h. Upon completion, the reaction was quenched with water (15.0 mL) and stirred for 15 min. The organic layer was separated and aqueous layer was washed with dichloromethane (2 × 10 mL). Combined organic layers were dried over sodium sulfate and evaporated to obtain a yellow oil. The crude product was purified by column chromatography using silica gel and 2% methanol in chloroform as the eluent system to obtain **7** as a yellow solid (808 mg, 96%), mp 120-125 °C, R_f: 0.76 (9.5:0.5 v/v, $CHCl_3$:CH_3OH); IR: υ 2976, 1712, 1663, 1554, 1444, 1123, 1019, 708 cm^{-1}; ^1H NMR ($CDCl_3$) δ 8.84 (s, 1H), 7.88-7.86 (d, 2H, *J*=8.0 Hz), 7.69 (s, 1H), 7.56-7.47 (m, 3H), 4.40 (q, 2H, *J*=8.0 Hz), 4.05 (s, 3H), 1.42 (t, 3H, *J*=8.0 Hz); LRMS (ES^+) *m/z* = 274 ($[M+H]^+$, 100%).

4-Benzamido-1-methylimidazole-2-carboxylic acid (8). To the slurry of ethyl 4-benzamido-1-methylimidazole-2-carboxylate (**7**, 800 mg, 2.9 mmol) in methanol:water (1:1, 20.0 mL) was added aqueous sodium hydroxide solution (2.0 M, 1.8 mL). The reaction mixture was refluxed for 45 min and then distilled to remove methanol. It was then made acidic (pH = 3-4) by adding dilute hydrochloric acid (6.0 M) to obtain 4-benzamido-1-methylimidazole-2-carboxylic acid (**8**) as a white solid which was filtered, washed with water and dried in vacuum at room temperature (664 mg, 85%), mp 156-158 °C, R_f: 0.24 (3:7 v/v, $CHCl_3$:CH_3OH); IR: υ 3445, 2254, 2127, 1661, 1033, 818, 753 cm^{-1}; ^1H NMR ($CDCl_3$) δ 11.05 (s, br, 1H), 8.04-8.06 (d, 2H, *J*=8.0 Hz), 7.70 (s, 1H), 7.54-7.46 (m, 3H), 3.95 (s, 3H); LRMS (ES^+) *m/z* = 246 ($[M+H]^+$, 100%).

Monoamino Ph-ImPyIm polyamide (3). 4-Benzamido-1-methylimidazole-2-carboxylic acid (**8**, 107 mg, 0.44 mmol) was dissolved in DMF (2.0 mL). To the clear solution was added PyBOP (230 mg, 0.46 mmol) followed by diisopropylethylamine (DIPEA, 0.15 mL, 0.87 mmol). The reaction mixture was stirred at room temperature for 30 min and to it was added solution of **9**[18] (133 mg, 0.4 mmol) in DMF (1.0 mL). The reaction mixture was stirred for 3 days at room temperature under argon atmosphere and then poured on water (10.0 mL). It was extracted with chloroform (3 × 10.0 mL) and combined organic layers were dried over sodium sulphate and evaporated on rotary evaporator. The crude product was purified by column chromatography using silica gel and 10% methanol in chloroform to obtain monoamino Ph-ImPyIm polyamide (**3**) as a yellow solid (87 mg, 38%), mp 163-165 °C, R_f: 0.53 (1:1 v/v, $CHCl_3$:CH_3OH); IR: υ 2928, 1663, 1531, 1468, 1438, 1258, 1117, 907 cm^{-1}; ^1H NMR ($CDCl_3$) δ 8.83 (s, br, 1H), 8.48 (s, br, 1H), 8.09 (s, br, 1H), 7.94-7.92 (d, 2H, *J*=8.0 Hz), 7.62 (s, 1H), 7.55 (s, 1H), 7.53-7.51 (m, 3H), 7.42 (s, 1H), 7.29 (s, 1H), 6.76 (s, 1H), 4.49 (q, 2H, *J*=8.0 Hz), 4.12 (s, 3H), 4.03 (s, 3H), 3.98 (s, 3H), 2.55 (t, 2H, *J*=8.0 Hz), 2.32 (s, 6H); LRMS (ES^+) *m/z* = 561 ($[M+H]^+$, 100%); HRMS $[M+H]^+$ calculated for *m/z* $C_{27}H_{32}N_{10}O_4$ 561.2691; found *m/z* 561.2686.

Ph-ImPy(C_3H_6Cl)Im (11). 4-Benzamido-1-methylimidazole-2-carboxylic acid (**8**, 126 mg, 0.5 mmol) was dissolved in DMF (2.0 mL). To the clear solution was added PyBOP (280 mg, 0.55 mmol) followed by diisopropylethylamine (DIPEA, 0.18 mL, 1.0 mmol). The reaction mixture was stirred at room temperature for 30 min and to it was added solution of **10**[16] (208 mg, 0.5 mmol) in DMF (1.0 mL). The reaction mixture was stirred for 3 days at room temperature under argon atmosphere and then poured on water (10.0 mL). It was extracted with chloroform (3 × 10.0 mL) and combined organic layers were dried over sodium sulphate and evaporated on rotary evaporator. The crude product was purified by column chromatography using silica gel

and 15% methanol in chloroform to obtain Ph-ImPy(C_3H_6Cl)Im (**11**) as a white solid (104 mg, 34%), mp 238-240 °C, R_f: 0.44 (1:1 v/v, $CHCl_3$:CH_3OH); IR: υ 3300, 1654, 1450, 1111, 1030, 930 cm^{-1}; ^1H NMR ($CDCl_3$) δ 8.87 (s, br, 1H), 8.49 (s, br, 1H), 8.25 (s, br, 1H), 8.05 (s, br, 1H), 7.94-7.92 (d, 2H, *J*=8.0 Hz), 7.62 (s, 1H), 7.59-7.55 (m, 3H), 7.52 (s, 1H), 7.45 (s, 1H), 7.40 (s, 1H), 4.53 (t, 2H, *J*=8.0 Hz), 4.13 (s, 3H), 4.03 (s, 3H), 3.54-3.48 (m, 4H), 2.57 (t, 2H, *J*=8.0 Hz), 2.33 (s, 6H), 2.31 (t, 2H, *J*=8.0 Hz); LRMS (ES^+) *m/z* = 623 ($[M+H]^+$, 50%).

Diamino Ph-ImPy*Im polyamide (5). The solution of Ph-ImPy(C_3H_6Cl)Im (**11**, 10 mg, 0.016 mmol) in dry methanol (5.0 mL) was purged with dry ammonia for 6.0 h in a re-sealable tube. The tube was sealed tightly and heated at 80 °C for 12 h. After completion of reaction, the solution was evaporated on a rotary evaporator to dryness to obtain the crude product. Purification by column chromatography (silica, gradient, 0:100-100:0%v/v, $CHCl_3$/MeOH) gave the desired diamino polyamide **5** as a white solid (6.0 mg, 62%), mp 150-152 °C, R_f: 0.41 (15:5:1% v/v, $CHCl_3$:CH_3OH:NH_3); IR: υ 3421, 3359, 3310, 3080, 3000, 2973, 1651, 1541, 1471, 1405, 1257, 1188, 1084, 1012, 847 cm^{-1}; ^1H NMR (CD_3OD) δ 7.98 (d, 1H), 7.57-7.64 (m, 6H), 7.49 (s, 1H), 7.07 (s, 1H), 4.52 (t, *J*=4.0 Hz, 2H), 4.13 (s, 3H), 4.05 (s, 3H), 3.52 (t, *J* =4.0 Hz, 2H), 2.81 (t, *J*=4.0 Hz, 2H), 2.60 (t, *J* =8.0 Hz, 2H), 2.35 (s, 6H), 2.01 (t, *J* =8.0 Hz, 2H), 2.11 (quint, *J* =4.0 Hz, 2H); LRMS (ES^+) *m/z* (rel intensity) 601 ($[M+H]^+$, 90%), 301 (100%); HRMS $[M+H]^+$ calculated for $C_{29}H_{37}N_{11}O_4$ *m/z* 601.2635; found *m/z* 601.2634.

Buffers. 10 mM phosphate buffers were prepared at the following salt (NaCl) concentrations: 12.5 mM ($PO_4$0), 50 mM ($PO_4$5), 200 mM ($PO_4$20) with the addition of 1 mM EDTA, pH 6.2.

DNase I footprinting. A radiolabeled DNA fragment of 125 base pairs was generated by polymerase chain reaction as described previously [21] . The resulting labeled fragment was purified on a Bio-Gel P-6 column (Bio-Rad) followed by agarose gel electrophoresis and isolated using a Mermaid kit (MP biomedicals) according to the manufacturer's instructions. DNase I digestions were conducted in a total volume of 8 μL. In each case, the labeled DNA fragment (2 μL, 200 counts s^{-1}) was incubated for 30 min at room temperature in 4 μL of TN binding buffer (10 mM Tris Base, 10 mM NaCl, pH 7) containing the required drug concentration. Cleavage by DNase I was initiated by addition of 2 μL of DNase I solution (20 mM NaCl, 2 mM $MgCl_2$, 2 mM $MnCl_2$, DNase I 0.02U, pH 8.0) and stopped after 3 min by snap freezing the samples on dry ice.

The nuclease-digested samples were subsequently lyophilized to dryness and resuspended in 5 μL of formamide loading dye (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol). Following heat denaturation for 5 min at 90 °C, the samples were loaded on a denaturing polyacrylamide (10%) gel (Sequagel, National Diagnostics, UK) containing urea (7.5 mM). Electrophoresis was carried out for 2 h at 1650 V (~70 W, 50 °C) in 1× TBE buffer. The gel was then transferred onto Whatman 3MM and dried under a vacuum at 80 °C for 2 h. The gel was exposed overnight to Fuji medical X-Ray film and developed on a Konica Medical Film Processor SRX-101A.

Thermal denaturation (T_M). Thermal denaturation data were obtained using the Cary 100 BioMelt (Varian) spectrophotometer with DNA (1 μM) in $PO_4$0 and the required compound (3 μM), using the procedure previously reported.[7a] Thermal melts were obtained for each compound using 5'-ACGCGT-3', 5'-AAATTT-3' and 5'-ACACTT-3'.

Circular dichroism (CD). CD studies were performed using the Olis DSM20 instrument. Each run was performed over 400–220 nm wavelength range (180 increments) and an integration time of 1 s and the average of two scans were used for analysis. The required compound (500 μM in double distilled H_2O) was titrated in 1 molar equivalents into the required DNA (160 μL of 9 μM DNA), in $PO_4$5, until saturation was observed. Data analysis was performed as previously described.[7a] CD experiments were performed with 5'-ACGCGT-3', 5'-AAATTT-3' and 5'-ACACTT-3'.

Surface plasmon resonance (SPR). SPR measurements were performed with a four-channel BIAcore T100 optical biosensor system (Biacore, GE Healthcare Inc.). 50-Biotin-labeled DNA hairpin duplex samples were immobilized onto streptavidin-coated sensor chips (BIAcore SA) as previously described.[20] Three flow cells were used to immobilize the DNA oligomer samples, while a fourth cell was left blank as a control. The SPR binding experiments were performed in 0.01M cacodylic acid (CCA) solution at pH 6.25 containing 0.001 M EDTA (disodium ethylenediamine tetraacetate), 0.1 M NaCl, and 0.005% v/v surfactant P20 (Biacore AB) was used to reduce the nonspecific binding of polyamides to the fluidics and sensor chip surface. A 0.01 M *N*-[2-hydroxyethyl]piperazine-N0-[2-ethanesulfonic acid] (HEPES) solution at pH 7.4 containing 0.15 M NaCl, 0.003 M EDTA, and 0.005% v/v surfactant P20 was used during the DNA immobilization process. All buffers were degassed and filtered prior to experiments. DNA sequences were obtained from Integrated DNA Technologies (San Diego, CA) with HPLC

purification and were used without further purification. The lyophilized 5'-biotin-labeled DNA hairpin constructs were dissolved in the appropriate amount of DI H_2O to create 1.0 mM DNA stock solutions. Further dilutions were made using 0.01 M HEPES buffer during the DNA immobilization step. DNA concentrations were determined spectrophotometrically using molar absorptivity coefficients calculated for each individual DNA sequence by using the nearest neighboring method for the singled-stranded DNA method. Stock solutions of **3** and **5** were prepared by dissolving the solid compound in the necessary amount of distilled H_2O to create a stock solution with a concentration of approximately 1.0×10^{-3} M. Stock solutions were kept frozen at 4 °C until experimental use to minimize any degradation of the compound. Samples for SPR experiments were prepared by a series of dilutions from the stock solution using a 0.01 M cacodylic acid buffer solution. The SPR experiments were conducted using a four-flow cell BIAcore 2000 biosensor instrument (GE Life Sciences). DNA hairpin constructs labeled with biotin at the 5' end were immobilized onto a streptavidin-coated sensor chip (sensor chip SA) as previously reported.[20a,b,22] The 5'-biotin-labeled DNA hairpins were immobilized on three of the four flow cells. The fourth was left blank and used as a control. All SPR experiments were performed at 25 °C and used 0.01 M CCA as the running buffer. The amount of DNA immobilized was approximately 400 response units (RU). This was achieved by continuously injecting ~20 μL of an approximately 50 nM DNA solution at a rate of 2 μL/min onto the sensor chip surface until a relative response of 400 units was reached. Binding data were obtained by injecting known concentrations and were analyzed with one or two site binding models as previously described using the following equation: [20a,b,22]

$$\boldsymbol{r} = (K_1 \times (C_{\text{free}} + 2 \times K_1 \times K_2 \times C^2_{\text{free}})) /(1 + K_1 \times C_{\text{free}} + K_1 \times K_2 \times C^2_{\text{free}}) \qquad \text{eq. 1}$$

where, r represents the moles of bound compound per mole of DNA hairpin duplex, K_1 and K_2 are macroscopic binding constants, and C_{free} is the free compound concentration in equilibrium with the complex.

Isothermal titration calorimetry (ITC). ITC analysis was performed using a VP-ITC microcalorimeter (MicroCal). Compound **3** was dissolved in 0.01 M cacodylic acid (CCA) buffer and the instrument equilibrated to 25 °C. ITC experiments were conducted using the "model-free" method as previously reported.[23] After an initial delay of 300 s, compound **3** (80 µM) was titrated, via 30 injections (10 µL for 20 s, repeated every 300 s), into 30 µM DNA (0.01 M CCA). The data were imported to Origin 7.0, which was used to integrate the area under the curve as a function of time. In order to normalize for nonspecific heat components, including heats of dilution, nonspecific heat component integrations were subtracted from the reaction integrations [5]. A linear fit of the normalized ΔH (y-axis) values versus mol ratio (x-axis) was done using KaleidaGraph 4.0. Extrapolation back to 0 mol ratio of this plot yields the enthalpy (ΔH) of the reaction. The DNA used in this experiment was 5'-ACGCGT-3', 5'-AAATTT-3', and 5'-ACCGGT-3'.

4.7 Acknowledgments

The authors thank the NSF [CHE 0809162, CHE 0550992, and CHE 0922623 (NMR)], Cancer Research UK (C2259/A9994 to JAH) and the Georgia Research Alliance for their generous support.

4.8 References

1. Mitra, S. N.; Wahl, M. C.; Sundaralingam, M. *Acta Crystallogr. Sect. D* **1999**, *55*, 602; (b) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci.* **1989**, *86*, 5723; (c) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc*. **1990**, *112*, 1393.
2. (a) Chou, C. J.; Farkas, M. E.; Tsai, S. M.; Alvarez, D.; Dervan, P. B.; Gottesfeld, J. M. *Mol. Cancer Ther.* **2008**, *7*, 769; (b) Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Nat. Struct. Biol.* **1996**, *3*, 321; (c) Chenoweth, D. M.; Dervan, P. B. *Proc. Natl. Acad. Sci.* **2009,** *106* 13175.
3. (a) Hochhauser, D.; Kotecha, M.; O'Hare, C.; Morris, P. J.; Hartley, J. M.; Taherbhai, Z.; Harris, D.; Forni, C.; Mantovani, R.; Lee, M.; Hartley, J. A. *Mol. Cancer Ther.* **2007**, *6*, 346; (b) Dwyer, T. J.; Geierstanger, B. H.; Bathini, Y.; Lown, J. W.; Wemmer, D. E. *J. Am. Chem. Soc.* **1992,** *114*, 5911.
4. (a) Shinohara, K.; Bando, T.; Sugiyama, H. *Anticancer Drugs* **2010**, *21*, 228; (b) Minoshima, M.; Bando, T.; Shinohara, K.; Sugiyama, H. *Nucleic Acids Symp. Ser.* **2009**, *23*, 69; (c) Borman, S. *Chem. Eng. News* **2010**, *88*, 50.
5. Buchmueller, K. L.; Bailey, S. L.; Matthews, D. A.; Taherbhai, Z. T.; Register, J. K.; Davis, Z. S.; Bruce, C. D.; O'Hare, C.; Hartley, J. A.; Lee, M. *Biochemistry* **2006**, *45*, 13551.
6. (a) Verma, R.; Patapoutian, A.; Gordon, C. B.; Campbell, J. L. *Proc. Natl. Acad. Sci.* **1991**, *88*, 7155; (b) Wu, X.; Lee, H. *Oncogene* **2002**, *21*, 7786; (c) Yamada, M.; Sato, N.; Taniyama, C.; Ohtani, K.; Drong, R. F.; Weiland, K. L.; Slightom, J. L.; Sclafani, R. A.; Hollingsworth, R. E. *Gene* **1998**, *28*, 133.
7. (a) Lacy, E. R.; Le, N. M.; Price, C. A.; Lee, M.; Wilson, W. D. *J. Am. Chem. Soc.* **2002**, *124*, 2153; (b) Buchmueller, K. L.; Staples, A. M.; Howard, C. M.; Horick, S. M.; Uthe, P. B.; Le, N. M.; Cox, K. K.; Nguyen, B.; Pacheco, K. A. O.; Wilson, W. D.; Lee, M. *J. Am. Chem. Soc.* **2005**, *127*, 742; (c) Buchmueller, K. L.; Staples, A. M.; Uthe, P. B.; Howard, C. M.; Pacheco, K. A. O.; Cox, K. K.; Henry, J. A.; Bailey, S. L.; Horick, S. M.; Nguyen, B.; Wilson, W. D.; Lee, M. *Nucleic Acids Res*. **2005**, *33*, 912; (d) Lee, M.; Rhodes, A. L.; Wyatt, M. D.; Forrow, S.; Hartley, J. A. *Biochemistry* **1993**, *32*, 4237.
8. Brown, T.; Mackay, H.; Turlington, M.; Sutterfield, A.; Smith, T.; Sielaff, A.; Westrate, L.; Bruce, C.; Kluza, J.; O'Hare, C.; Nguyen, B.; Wilson, D. W.; Hartley, J. A.; Lee, M. *Bioorg. Med. Chem.* **2008**, *16*, 5266.
9. Collar, C. J.; Lee, M; Wilson, W. D. *J. Chem. Inf. Model.* **2010**, *50*, (9), 1611.
10. Thomas, M.; Varshney, U.; Bhattacharya, S. *Eur. J. Org. Chem*. **2002**, 3604.

11. Lajiness, J.; Sielaff, A.; Mackay, H.; Brown, T.; Kluza, J.; Nguyen, B.; Wilson, W. D.; Lee, M.; Hartley, J. A. *Med. Chem.* **2009**, *5*, (3), 216.
12. Lee, M.; Lori, A. W.; Nobles, J. A.; Forrow, S. M.; Hartley, J. A. *Bioorg. Med. Chem. Lett.* **1991**, *1*, 595.
13. (a) Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007**, *35*, 363; (b) Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res.* **2004**, *32*, 2802; (c) Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci.* **2003**, *100*, 12063.
14. Lown, J. W.; Krowicki, K.; Bhat, U. G.; Skorobogaty, A.; Ward, B.; Dabrowiak, J. C. *Biochemistry* **1986**, *25*, 7408.
15. Satz, A. L.; Bruice, T. C. *Acc. Chem. Res.* **2002**, *35*, 86.
16. Babu, B.; Liu, Y.; Plaunt, A.; Riddering, C.; Ogilvie, R.; Westrate, L.; Davis, R.; Ferguson, A.; Mackay, H.; Rice, T.; Chavda, S.; Wilson, D.; Lin, S.; Kiakos, K.; Hartley, J. A.; Lee, M. *Biochem. Biophys. Res. Commun.* **2011**, *404*, 848.
17. Lee, M.; Rhodes, A.; Wyatt, M. D.; Forrow, S.; Hartley, J. A. *Biochemistry* **1993**, *32*, 4237.
18. Chavda, S.; Liu, Y.; Babu, B.; Davis, R.; Sielaff, A.; Ruprich, J.; Westrate, L.; Tronrud, C.; Ferguson, A.; Franks, A.; Tzou, S.; Adkins, C.; Rice, T.; Mackay, H.; Kluza, J.; Tahir, S. A.; Lin, S.; Kiakos, K.; Bruce, C. D.; Wilson, D. W.; Hartley, J. A.; Lee M. *Biochemistry* **2011**, *50*, 3127.
19. (a) Westrate, L.; Mackay, H.; Brown, T.; Nguyen, B.; Kluza, J.; Wilson, D. W.; Lee, M.; Hartley, J. A. *Biochemistry* **2009**, *48*, 5679; (b) Lown, J. W.; Krowicki, K.; Bhat, U. G.; Skorobogaty, A.; Ward, B.; Dabrowiak, J. C. *Biochemistry* **1986**, *25*, 7408.
20. (a) Liu, Y.; Wilson, W. D. *Methods Mol. Biol.* **2010**, 613, 1; (b) Lacy, E. R.; Minh Le, N.; Price, C. A.; Lee, M.; Wilson., W. D. *J. Am. Chem. Soc.* **2002**, *124*, 2153; (c) Minh Le, N.; Sielaff, A.; Cooper, A. J.; Mackay, H.; Brown, T.; Kotecha, M.; O'Hare., C.; Hochauser, D.; Lee, M.; Hartley, J. A. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 6161; (d) Buchmueller, K. L.; Staples, A. M; Uthe, P. B.; Howard, C. M.; Pacheco, K. A. O.; Cox, K. K.; Henry, J. A.; Bailey, S. L.; Horick, S. M.; Nguyen, B.; Wilson, D. W.; Lee, M. *Nucleic Acids Res.* **2005**, *33*, 912; (e) Brown, T.; Taherbhai, Z.; Sexton, J.; Sutterfield, A.; Turlington, M.; Jones, J.; Stallings, L.; Stewart, M.; Buchmueller, K.; Mackay, H.; O'Hare, C.; Kluza, J.; Nguyen, B.; Wilson, D. W.; Lee, M.; Hartley, J. A. *Bioorg. Med. Chem.* **2007**, *15*, 474; (f) Buchmueller, K.; Staples, A. M.; Uthe, P. B.; Howard, C. M.; Pacheco, K. A. O.; Cox, K. K.; Henry, J. A.; Bailey, S. L.; Horick, S. M.; Nguyen, B.; Wilson, D. W.; Lee, M. *J. Am. Chem. Soc.* **2005**, *127*, 742.
21. O'Hare, C. C.; Uthe, P.; Mackay, H.; Blackmon, K.; Jones, J.; Brown, T.; Nguyen, B.; Wilson, W. D.; Lee, M.; Hartley, J. A. *Biochemistry* **2007**, 46, 11661.
22. Lacy, E. R.; Nguyen, B.; Minh Le, N.; Cox, K. K.; O' Hare, C.; Hartley, J. A.; Lee, M.; Wilson, D. *Nucleic Acids Res.* **2002**, *30*, 1834.
23. Ren, J.; Jenkins, T. C.; Chaires, J B. *Biochemistry* **2000**, *39*, 8439-8447.

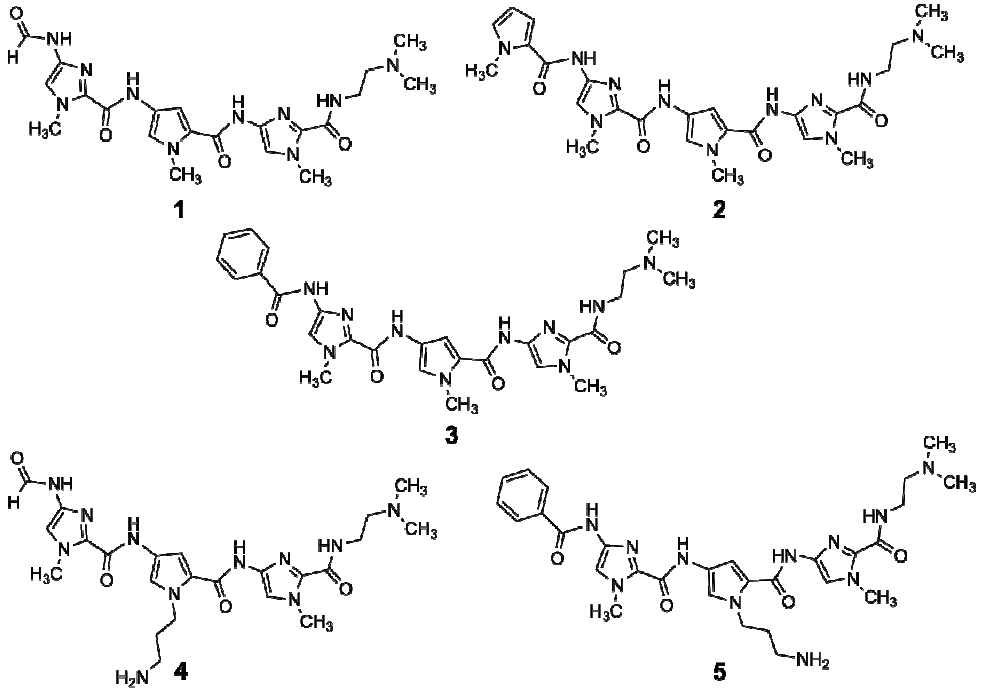


Figure 4.1 Structures of f-ImPyIm (**1**), tetraamide polyamide PyImPyIm (**2**), monoamino polyamide, Ph-ImPyIm (**3**), an orthogonally positioned diamino polyamide f-ImPy*Im (**4**) and an orthogonally positioned diamino polyamide Ph-ImPy*Im (**5**), Py* represents the *N*-(3-aminopropyl)-4-amidopyrrole-2-carboxamido moiety.

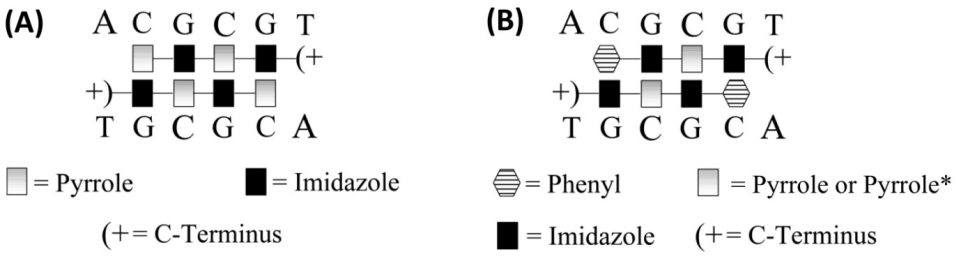


Figure 4.2 Schematic of PyImPyIm (**4**) (**A**) and phenyl containing polyamides Ph-ImPyIm (**3**) and Ph-ImPy*Im (**5**) (**B**) binding to the cognate DNA sequence 5'-ACGCGT-3'.

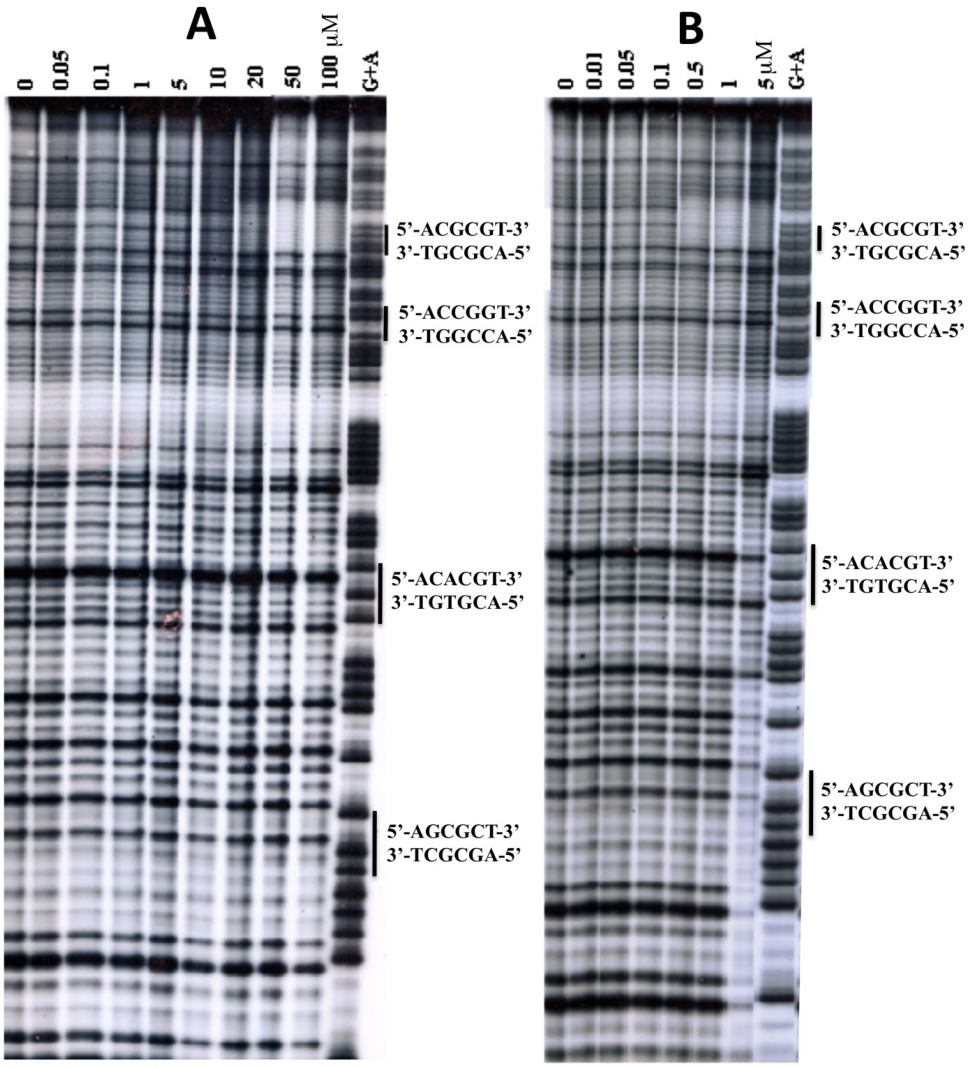


Figure 4.3 DNase I footprinting of monoamino polyamide Ph-ImPyIm (**3**) (**A**) and diamino polyamide Ph-ImPy*Im (**5**) (**B**).

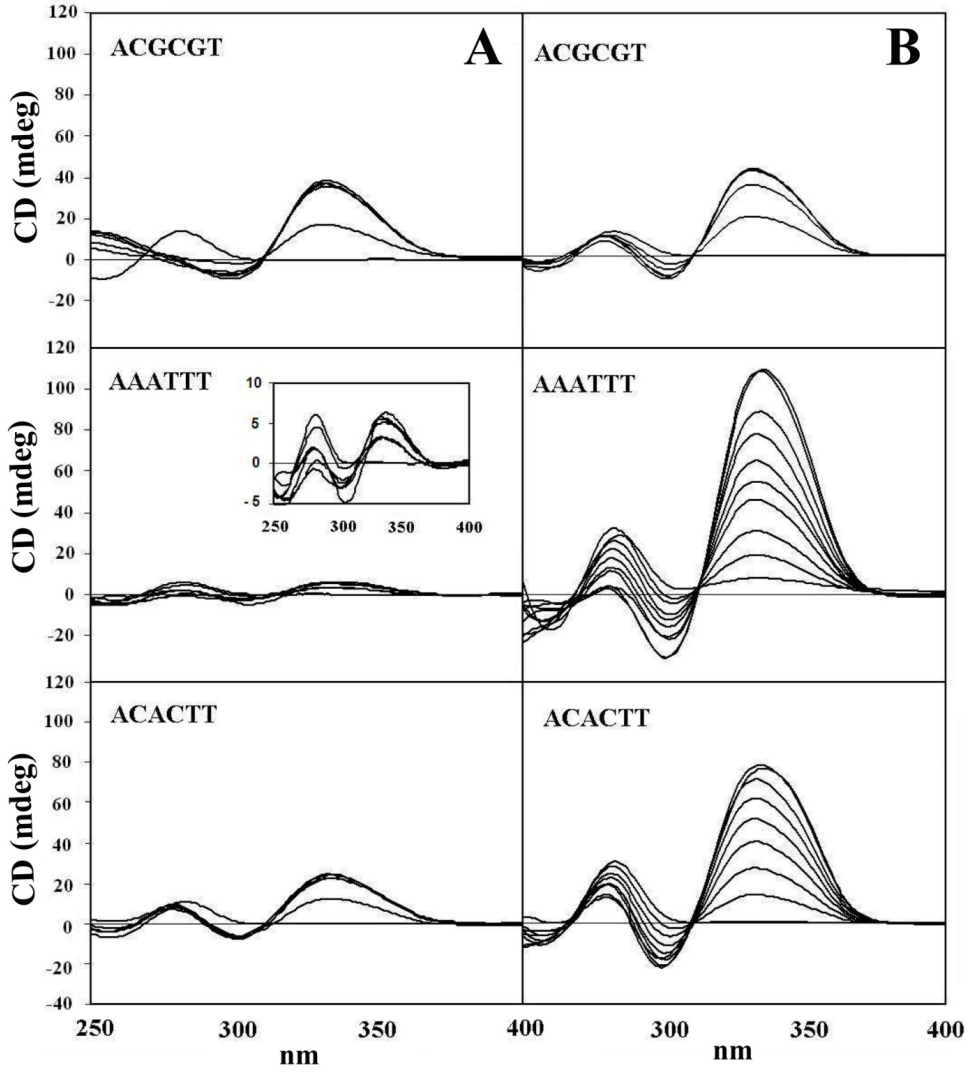


Figure 4.4 CD data for Ph-ImPyIm (**3**) (**A**) and its diamino counterpart **5** (**B**) with 5'-ACGCGT-3', 5'-AAATTT-3' and 5'-ACACTT-3'. Data were obtained using $PO_4$5 buffer.

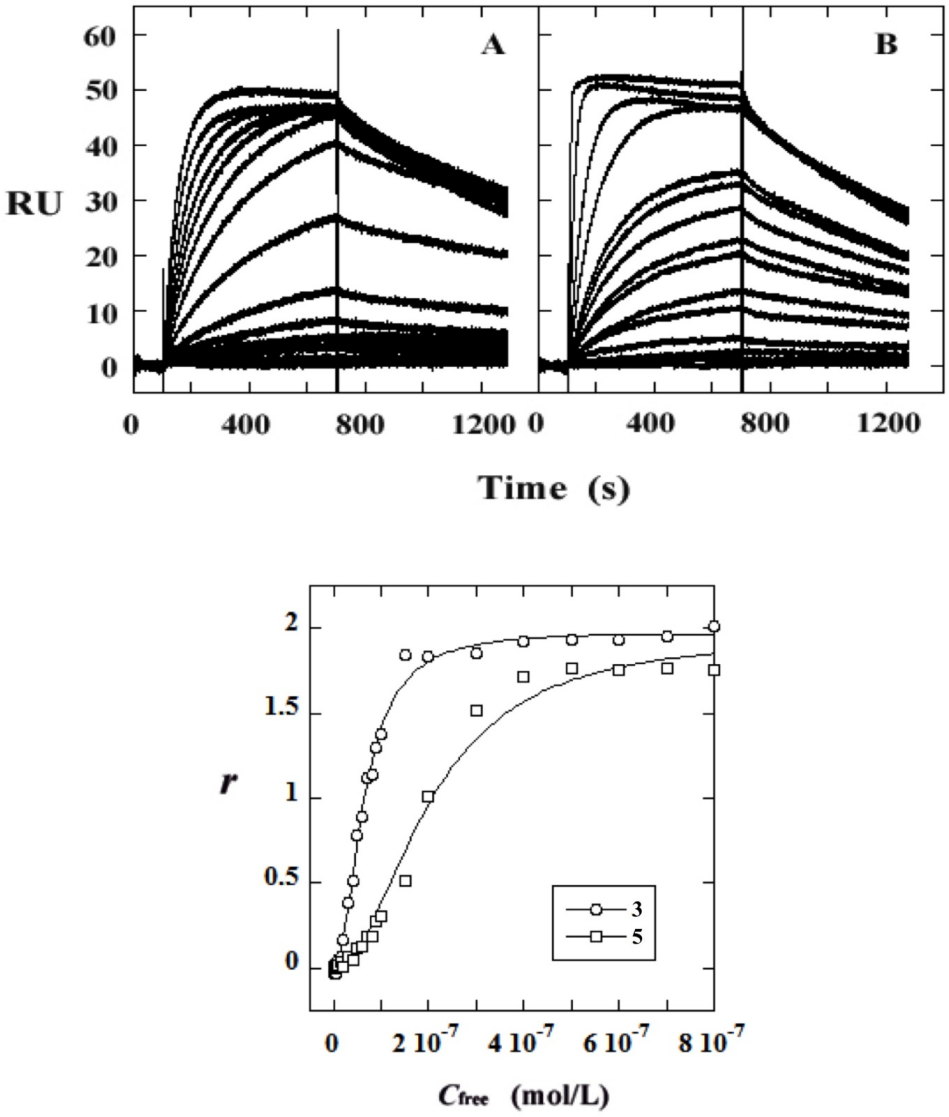


Figure 4.5 SPR sensorgrams (**A**) and (**B**) showing binding of **3** and **5** to 5'-ACGCGT-3', respectively, at concentration range from 0 to 0.8 μM. (**C**) Plot of *r vs* concentration including best fit lines.

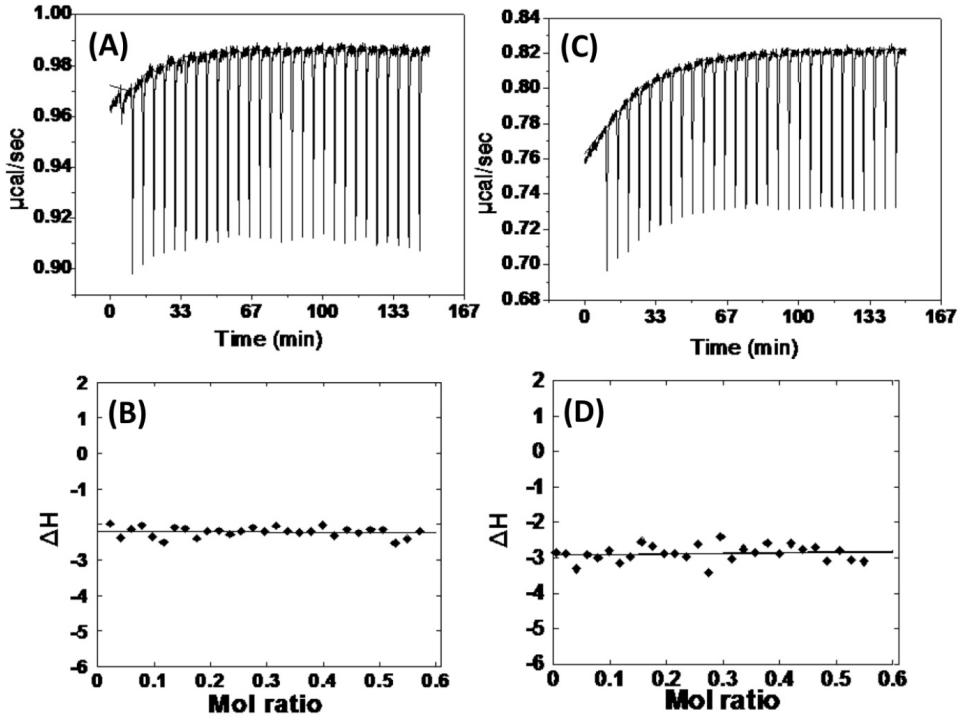
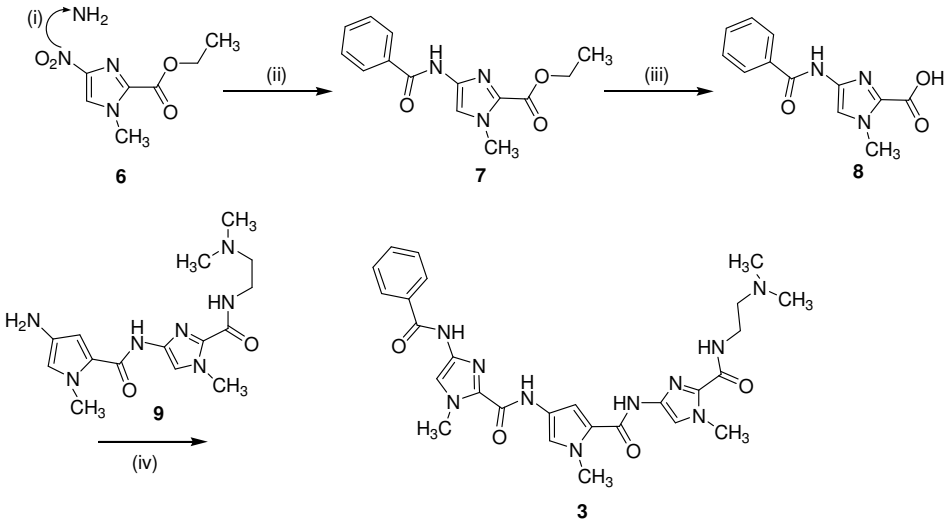
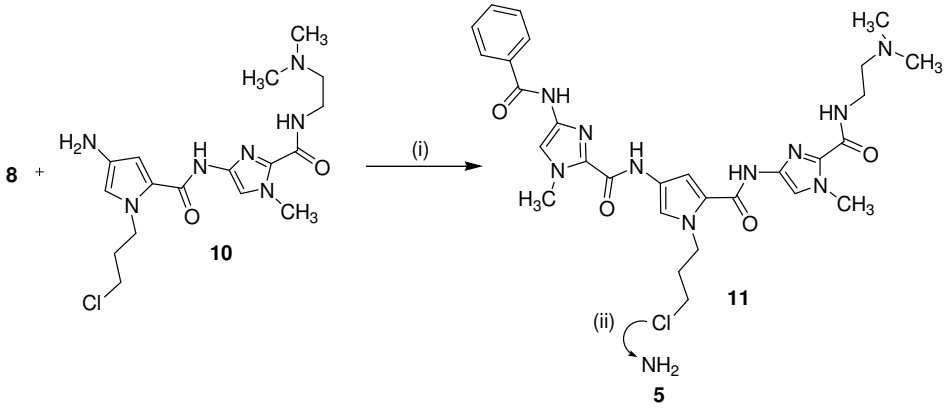


Figure 4.6 (**A**) Raw ITC data for the "model free" titration of **3** into 5'-ACGCGT-3'.[a] (**B**) The integrated heats for each injection of **3** used for the titration (red dots) and a linear fit (red line) of that data.[b] (**C**) Raw ITC data for the "model free" titration of **5** into 5'-ACGCGT-3'.[a] (**D**) The integrated heats for each injection of **5** used for the titration (red dots) and a linear fit (red line) of that data.[b] The titration consisted of 30-10 µL injections of 80 µM of **3** or **5** into 30 µM DNA. The titration was completed up to a mole ratio of ~0.6:1 (ligand :DNA). [b] The linear fits were completed using KaleidaGraph 4.0.



Scheme 4.1 Reagents and conditions (i) H_2, 5% Pd/C, ethanol, rt, 12 h; (ii) benzoyl chloride, TEA, DCM, rt, 12 h; (iii) NaOH, CH_3OH:H_2O (1:1), reflux, 45 min; (iv) PyBOP, DMF, DIPEA, rt, 3 days.



Scheme 4.2 Reagents and conditions (i) PyBOP, DMF, DIPEA, rt, 3 days; (ii) dry NH_3, methanol, 80 °C, 12 h.

Table 4.1 Thermal denaturation data and DNase I footprinting data for polyamides **3**, **5** and **1**

	ΔT_m (°C)			
Polyamide	**5'-ACGCGT-3'**	**5'-A_3T_3-3'**	**5'-ACACTT-3'**	**Footprinting (μM)[a]**
3	9	0	0	20
5	>25	7.6	0	0.5
1[a]	9	1	ND[b]	0.05

[a]Data taken from reference 5. [b]Not determined.

Table 4.2 Binding constants of polyamides **3**, **5** and **1** to DNA sequences 5'-ACGCGT-3' and5'-AAATTT-3'.

	Sequences	
Polyamide	**5'-ACGCGT-3'**	**5'-A_3T_3-3'**
3	$K_1 = 6.1 \times 10^4$ M^{-1} $K_2 = 3.7 \times 10^8$ M^{-1} $K_{eq} = 4.8 \times 10^6$ M^{-1} ΔG = -9.1 kcal/mol ΔH = -2.2 kcal/mol TΔS = 6.9 kcal/mol	$K_1 = 2.9 \times 10^5$ M^{-1} $K_2 = 1.2 \times 10^7$ M^{-1} $K_{eq} = 2.1 \times 10^6$ M^{-1} ΔG = -8.6 kcal/mol ΔH = -2.2 kcal/mol TΔS = 6.9 kcal/mol
5	$K_1 = 8.7 \times 10^5$ M^{-1} $K_2 = 2.7 \times 10^8$ M^{-1} $K_{eq} = 1.5 \times 10^7$ M^{-1} ΔG = -9.8 kcal/mol ΔH = -2.7 kcal/mol TΔS = 7.1 kcal/mol	ND[d]
1[a]	$K_1 = 1.9 \times 10^8$ M^{-1} ΔG = -9.5 kcal/mol ΔH = -7.7 kcal/mol TΔS = 1.8 kcal/mol	$K_1 = 2.9 \times 10^5$ M^{-1} $K_2 = 1.2 \times 10^7$ M^{-1} $K_{eq} = 2.1 \times 10^6$ M^{-1} ΔG = -8.6 kcal/mol ΔH = -2.2 kcal/mol TΔS = 6.9 kcal/mol

[a]Binding constants were measured by SPR. [b]No response detected.
[c]Data taken from reference 5. [d]Not determined.

CHAPTER FIVE:

AFFINITY AND KINETIC MODULATION OF NOVEL SYNTHETIC POLYAMIDE DERIVATIVES

The work presented in this chapter is based on a paper *"Affinity and kinetic modulation of novel synthetic polyamide derivatives"* which has currently been submitted to *Biochemistry* for publication. I am the 1st author of this manuscript and also conducted all of the SPR and ITC experiments, data analyses and writings, along with Dr. Yang Liu and Dr. W. David Wilson. Synthetic methods and write ups are attributed to Dr. Moses Lee of Hope College and various members of the Lee Research Group. Dr. John Hartley from the UK Drug-DNA Interactions Research Group is credited with the DNase I footprinting work and writings.

AFFINITY AND KINETIC MODULATION OF NOVEL SYNTHETIC POLYAMIDE DERIVATIVES

Joseph P. Ramos[a], Balaji Babu[b], Sameer Chavda[b], Yang Liu[a], Adam Plaunt[b], Amanda Ferguson[b], Mia Savagian[b], Megan Lee[b], Samuel Tzou[b], Shicai Lin[c], Konstantinos Kiakos[c], Moses Lee, John A. Hartley[c], and W. David Wilson[a]

[a]Department of Chemistry, Georgia State University Atlanta, GA, 30303, USA; [b]Department of Chemistry and the Division of Natural and Applied Sciences, Hope College 49423, USA; [c]Cancer Research, UK Drug-DNA Interactions Research Group, UCL Cancer Institute, Paul O' Gorman Building, 72 Huntley Street, London WCIE 6BT, UK

5.1 Abstract

Synthetic *N*-methyl imidazole and *N*-pyrrole containing polyamides that can form "stacked" dimers can be programmed to target and bind to specific DNA sequences and control gene expression. In order to accomplish this goal, the development of polyamides with lower molecular mass which allows for the molecules to rapidly penetrate cells and localize in the nucleus, along with increased water solubility, while maintaining DNA binding sequence specificity and high binding affinity, is key. In order to meet these challenges, six novel f-ImPy*Im polyamide derivatives that contain different orthogonally positioned moieties were designed to target 5'-ACGCGT-3'. The synthesis and biophysical characterization of six f-ImPy*Im were determined by CD, ΔT_M, DNase I footprinting, SPR, and ITC studies, and were compared with those of its parent compound, f-ImPyIm. The results gave evidence for the minor groove binding and selectivity of polyamides **1** and **6** for the cognate sequence 5'-ACGCGT-3', and with strong affinity $K_{eq} = 2.8 \times 10^8\ M^{-1}$ and $K_{eq} = 6.2 \times 10^7\ M^{-1}$, respectively. The six novel polyamides presented in this study demonstrated increased water solubility, while maintaining low molecular mass, sequence specificity, and binding affinity, addressing key issues in therapeutic development.

Keywords: Polyamides, diamino, dicationic, gene control, ACGCGT, MCB, *Mlu1*, cell-cycle box, pyrrole, imidazole, DNA, minor groove binder, sequence selectivity, binding affinity, kinetics

5.2 Introduction

Synthetic pyrrole (Py) and imidazole (Im) containing analogs (Figure 5.1) of the naturally occurring polyamide compounds distamycin and netropsin have been shown to selectively bind as a "stacked" dimer to the minor groove of DNA with high affinity and selectivity (*1-3*). The addition of a formamido (f) group to the N-terminus of polyamides (PA) increases the binding affinity and causes the stacking to be in a "staggered" position as opposed to "overlapped" stacking (*4-8*). Because of their high affinity and sequence selectivity, polyamides have demonstrated potential as therapeutics. Based on their ability to target, alter and control gene expression, especially those genes associated with such diseases as cancer, polyamides are of interest for the design and development of new types of therapeutics (*9-16*).

The polyamide, f-ImPyIm (Figure 5.1), binds to the sequence 5'-ACGCGT-3' as a stacked, staggered homodimer with a (quite) high binding affinity, $K_{eq} > 10^8\ M^{-1}$, for a small molecule (*5, 6, 8, 17, 18*). The sequence 5'-ACGCGT-3' is of significance due to its occurrence in the core sequence of the Mlu1 cell-cycle box (MCB), a transcriptional element found in the promoter region of the human Dbf4 (huDbf4 or ASK, activator S-kinase) gene. High levels of Cdc7 (cyclin dependent) kinase have been implicated in the development of various cancers and Dbf4 is a regulatory subunit of this kinase (*19-23*).

Our approach to developing modified derivatives of polyamides, such as f-ImPyIm, as therapeutic candidates is to develop analogues that will facilitate an increase in binding affinity, sequence selectivity, solubility, and biological activity, including cellular uptake and nuclear localization. An attractive approach for the design of potentially improved analogues of f-ImPyIm is the addition of different functional groups at the *N*5-position of the center pyrrole

coupled with a detailed study of their DNA binding affinity and sequence selectivity. To initiate this investigation six different f-ImPy*Im analogs were synthesized (shown in Figure 5.1), where Py* is the modified heterocycle, and their binding properties were examined using surface plasmon resonance (SPR), isothermal titration calorimetry (ITC), DNase I footprinting, thermal denaturation, and circular dichroism (CD). All synthesized polyamides investigated in this study retain the same f-ImPy*Im core and, therefore, have the same 5'-ACGCGT-3' cognate sequence.

5.3 Experimental

5.3.1 Surface Plasmon Resonance

SPR measurements were performed with a four-channel Biacore T200 optical biosensor system (Biacore, GE Healthcare Inc.). 5'-Biotin-labeled DNA hairpin duplex samples were immobilized onto a streptavidin-coated sensor chip (Biacore SA) as previously described (*5-8, 17, 18*). All SPR experiments were performed at 25 °C and used 0.01 M cacodylic acid (CCA), 0.1 M NaCl, 0.001 M ethylenediaminetetraacetic (EDTA), 0.05% (by vol) P20 surfactant, and pH 6.25 as the running buffer. The amount of DNA immobilized was approximately 400 response units (RU). This was achieved by continuously injecting ~20 μL of an approximately 50 nM DNA solution at a rate of 2 μL/min onto the sensor chip surface until a relative response of 400 units was reached. Binding data were obtained by injecting known concentrations over each flow cell of the chip and were analyzed with a two site binding model as previously described (*24-26*): $r = (K_1 \times (C_{free} + 2 \times K_1 \times K_2 \times C^2_{free})) / (1 + K_1 \times C_{free} + K_1 \times K_2 \times C^2_{free})$ where r represents the moles of bound compound per mole of DNA hairpin duplex, K_1 and K_2 are macroscopic binding constants, and C_{free} is the free compound concentration in equilibrium with the complex. The stoichiometry of the reaction, n, can be determined directly from an SPR

experiment. This is best determined by an experiment in which the concentration of injected molecules (flow solution of ligand) flowed over each cell of the chip and this leads to no change in RU_{obs} when an increase in concentration occurs (*24, 26*). This is the steady-state RU value. For the binding of a small molecule, such as the polyamides presented in this study (PA), to a specific DNA site (NA) immobilized on the chip, to form a polyamide/DNA complex, the $RU_{max/pred}$ is dependent on the amount of immobilized DNA, RU_{NA}, the molecular weights of both the polyamide (M_{PA}) and the immobilized DNA component (M_{NA}), and the refractive index (n'), with the concentration (c) for polyamide $(dn'/d_c)_{PA}$ and DNA $(dn'/d_c)_{NA}$ given the equation:

$$RU_{max/pred} = RU_{NA} \times (M_{PA}/M_{NA}) \times [(dn'/d_c)_{PA}/(dn'/d_c)_{NA}] \qquad \text{Eq. 1}$$

RU_{obs}, the steady-state response, for a reaction depends on the fraction of DNA sites with compound bound, f_{bound}, for each concentration of ligand flow solution and the stoichiometry, n, given the equation:

$$RU_{obs} = RU_{max/pred} \times n \times f_{bound} \qquad \text{Eq. 2}$$

At the maximum, where f_{bound} is equal to 1, the stoichiometry, n, is given by:

$$n = (RU_{max})_{obs} / (RU_{max})_{pred} \qquad \text{Eq. 3}$$

K_1 and K_2, which are dependent on f_{bound}, are determined by plotting steady-state RU_{obs} values from eq. 2 versus concentration of free ligand (C_{free}) (*24, 26*), as shown in Figure 5.2B. The plot of $RU_{obs(ss)}$ versus C_{free} is fit using KaleidaGraph 4.0 software and a defined model for 2:1 binding based on the stoichiometry and is given on page 99. The binding process is considered to be cooperative when K_2 is at least four times greater than K_1.

5.3.2 Isothermal Titration Calorimetry (ITC)

ITC analysis was performed using a VP-ITC microcalorimeter (MicroCal, MA). The DNA and compound sample were prepared in the exact same [0.01 M cacodylic acid (CCA), 0.1 M NaCl, 0.001 M ethylenediaminetetraacetic acid (EDTA), pH 6.25] buffer solution. The instrument was equilibrated at a temperature of 25 °C (298 K). The ITC experiments were conducted using the "excess DNA" method ("model-free" method), where an excess of DNA is used and does not allow for the reaction to reach an equilibrium end point, as previously reported (*27-29*). Briefly, after an initial delay of 300 s, compounds (80 μM) were titrated via 30 injections (10 μL for 20 s, repeated every 300 s), into 30 μM DNA. All ITC experiments were repeated in triplicate and reproducibility noted. The data were analyzed using the same method as previously reported (*29*). Origin 7.0 was used and the area under each curve integrated as a function of time. A linear fit of the integrated heats was then employed using KaleidaGraph 4.0 and this was subtracted from the reaction integrations to normalize for nonspecific heat components. ΔG was calculated from $\Delta G = -RT \ln K_{eq}$, where, R is 1.987 cal mol^{-1} K^{-1} and T is measured in K.

5.3.3 DNase I Footprinting

DNase I footprinting experiments were conducted using a 130-bp 5'-[^{32}P]-radiolabeled DNA fragment to determine sequence specificity for polyamides **1-6**. The engineered DNA fragment contained each of the following sequences, the cognate sequence 5'-ACGCGT-3', as well as two other DNA sequences with the same CG composition in differing arrangements, 5'-AGCGCT-3' and 5'ACCGGT-3'. The protocol was reported previously (*6*).

5.3.4 Thermal Denaturation

Thermal denaturation studies were performed in duplicate using a Cary Bio 100 spectrophotometer UV/vis instrument (Palo Alto, CA) and cells with a 10 mm pathlength using a previously reported method (*6*). This method was used to support the sequence selectivity of polyamides **1-6** demonstrated by DNase I footprinting analyses (*6*). ΔT_M values were determined from the difference in melting temperatures of the DNA-polyamide complex and duplex DNA alone. For each experiment 1.0 μM of DNA and 3.0 μM of polyamide were used.

5.3.5 Circular Dichroism (CD)

CD spectroscopy was used to probe the binding of polyamides **1-6** in the minor groove of double-stranded DNA using an OLIS (Bogart, GA) DSM20 spectropolarimeter with a 10 mm path length cuvette and a band pass of 2.4 nm (*6*). These experiments were conducted in duplicate by titrating each polyamide with DNA solutions comprised of the DNA sequences tested in the thermal denaturation studies. In all cases, a fixed DNA concentration of 9 μM was used as were the ratios of polyamides (1, 2, 3, 4, 5, 6, 8, and 10 molar equivalents) titrated into the solution.

5.3.6 Synthesis

f-IP(C_3NH_2)I (1)

The synthesis of diamino f-IPI (**1**) was reported earlier (*22*). A more efficient synthetic approach is reported herein. A solution of f-IP(C_3Cl)I (**2**) tail (100 mg, 0.183 mmol) in dry MeOH (5 mL) in a sealed tube was purged with dry NH_3 (g) for 6 h. The tube was sealed tightly and heated at 60 ºC for 24 h. The solvent was removed by evaporation and the residue purified by flash column chromatography eluting in MeOH:$CHCl_3$ (0:100-100:0) to give diamino f-IPI

(**1**) as a yellow solid (30 mg, 31%). The ^{1}H-NMR spectrum is identical to the product obtained from hydrogenation of the azide precursor (*22*).

f-IP (C_3Cl)I (2)

Nitro-IP(C_3Cl)I (200 mg, 0.366 mmol) was reduced in the presence of hydrogen over 5% Pd/C (100 mg) in cold MeOH (100 mL) for 18 h. The reaction mixture was filtered over celite and the catalyst washed thoroughly with MeOH. The solvent was removed by evaporation and the residual MeOH removed by co-evaporation with dry DCM (3× 2 mL). The resulting product, a dark orange solid, was dried under high vacuum while being protected from light. Formic acid (1 ml, kept at 0 ºC prior to use) was added to acetic anhydride (2 ml) at 0 ºC and the solution heated at 50 ºC for 30 min. The solution was then added to an ice-cold solution of amine in dry DCM dropwise (~ 10 min.) and the solution stirred overnight. The excess anhydride was quenched with methanol (30 ml), and the solvent removed by evaporation. A basic aqueous work-up was performed and the residue purified by flash column chromatography eluting in MeOH:$CHCl_3$ (0:100-100:0) to give chloro f-IPI (**2**) as a yellow solid (113 mg, 56.3%); R_f: [MeOH:$CHCl_3$ (50:50)] 0.43; mp. 122 ºC; IR: υ (NaCl) cm^{-1} 3390, 3184, 2925, 2359, 1651, 1537, 1384, 1070; ^{1}H-NMR δ_H (400 MHz, $CDCl_3$) 8.11 (1H, s), 7.86 (2H, d, *J* 2), 7.40 (1H, s), 7.37 (2H, d, *J* 2), 4.54 (2H, t, *J* 6), 4.23 (3H, s), 3.57 (2H, q, *J* 7), 3.5 (2H, t, *J* 6.2), 2.56 (2H, s), 2.34 (2H, t, *J* 6), 2.18 (3H, s); LRMS (ES^+) m/z (rel intensity) 548 ([M+H]$^+$, 100%); HRMS [M+H$^+$, $C_{23}H_{31}ClN_{10}O_4^+$] calculated: 547. 2294. Found: 547.2286.

f-IP(C_3N_3)I (3)

To a solution of the chloro-polyamide fIP(C_3Cl)I (185 mg, 0.342 mmol) was added NaN_3 (103 mg, 1.58 mmol) in dry DMF (2 mL). The solution was heated at 80 °C overnight. The DMF

was removed by Kugelrohr distillation, a basic aqueous work-up performed and the residue purified by flash column chromatography eluting in MeOH:$CHCl_3$ (0:100-100:0) to give azido f-IPI (**3**) as a yellow solid (180 mg, 0.33 mmol, 95.8%); R_f: [MeOH:$CHCl_3$ (30:70)] 0.33; mp 112 ºC; IR: υ (NaCl) cm^{-1} 3329, 2921, 2359, 2099, 1659, 1531, 1470, 1402, 1082; NMR δ_H (400 MHz, $CDCl_3$) 8.85 (1H, s,), 8.37 (1H, s), 7.45 (1H, s), 7.39 (1H, s), 7.31 (1H, s), 6.85 (1H, s), 4.45 (2H, t, *J* 4.2 Hz), 4.08 (3H, s), 4.03 (3H, s), 3.49 (2H, t, *J* 4.2 Hz), 3.32 (2H, quart, *J* 4.2 Hz) 2.54 (2H, t, *J* 4.2 Hz), 2.32 (6H, s), 2.10 (2H, quint, *J* 4.2 Hz); LRMS (ES^+) m/z (rel intensity) 554 ($[M+H]^+$, 100%); HRMS [$M+H^+$, $C_{23}H_{32}N_{13}O_4^+$] calculated: 554.2700. Found: 554.2692.

f-IP(C_3alkyne)I (4)

NO_2P(C_3alkyne)I tail (100 mg, 0.241 mmol) was reduced in the presence of $SnCl_2$•$2H_2O$ (354 mg, 1.57 mmol) and 12 M HCl (1.1 mL) while refluxing in ethanol for 2 h. The reaction mixture was poured into water and the pH brought to 10 (2M NaOH). The solution was extracted with chloroform, the organic layers dried over anh. Na_2SO_4 and the solvent removed by evaporation. The corresponding amine was allowed to dry under high vacuum for 30 min. DiPEA (0.21 mL, 0.97 mmol) was added to a solution of 1-methyl-4-*N*-formamidoimidazole-2-carboxylic acid (*30*) (45 mg, 0.26 mmol) and PyBOP (138 mg, 0.265 mmol) in dry DMF under argon atmosphere DiPEA (0.21 mL, 0.97 mmol) and allowed to stir for 45 min. The dried amine was added to the activated acid in dry DMF over a period of 10 min. The solution was allowed to stir overnight. The DMF was removed by Kugelrohr distillation and the residue purified by flash column chromatography eluting in MeOH:$CHCl_3$ (0:100-100:0) to give alkynl f-IPI (**4**) as a light pink solid (38 mg, 24%); R_f: [MeOH:$CHCl_3$:liq NH_3 (5:15:1)] 0.31; mp 197-200 ºC; IR: υ (KBr) cm^{-1} 3373, 3285, 2950, 2113, 1669, 1663, 1539, 1472, 1405, 1243, 1123, 1018, 904, 752; ^{1}H-NMR δ_H (400 MHz, $CDCl_3$) 8.81 (1H, br s), 8.37 (1H, s), 8.18 (1H, br s), 7.58 (1H, s,), 7.46 (1H,

s), 7.37 (1H, s), 7.26 (1H, s), 4.59 (2H, t, *J* 6 Hz), 4.08 (3H, s), 4.02 (3H, s), 3.50 (2H, q, *J* 8.0 Hz), 2.56 (2H, t, *J* 6.0 Hz), 2.35 (6H, s), 2.30 (2H, t, *J* 5.6 Hz), 2.01 (3H, m); LRMS (ES^+) m/z (rel intensity) 571 ($[M+H]^+$, 75%); HRMS [$M+H^+$, $C_{25}H_{31}ClN_{10}O_4^+$] calculated: 571.2300. Found: 571.2297.

f-IP(C_3NHCOCH$_3$)I (5)

Acetyl chloride (24 µL, 0.034 mmol) was added dropwise to a solution of fIP(C_3NH_2)I tail (18 mg, 0.03 mmol) and dry Et_3N (4 µL, 0.03 mmol) in dry DCM (10 mL) and the solution stirred overnight. The solution was quenched with water and pH brought to 7 (1M NaOH). The solution was extracted with DCM (3 x 10 mL), the organic layers dried over anh. Na_2SO_4 and the solvent were removed by evaporation to yield a brown residue purified by flash column chromatography eluting in MeOH:$CHCl_3$ (0:100-100:0) to give acetamido f-IPI (**5**) as a light yellow solid (15 mg, 79%). R_f: [MeOH:$CHCl_3$:liq NH_3 (5:15:1)] 0.27; IR: υ (ATR) cm^{-1} 3482, 3257, 3079, 2952, 1654, 1630, 1551, 1534, 1455, 1438, 1306, 1257, 1211, 1185, 1167, 1151, 1126, 1090, 1053, 1039, 1013, 897, 864, 797, 775, 760, 670, 687, 656, 618, 575, 469, 447, 431, 415, 406; NMR δ_H (400 MHz, $CDCl_3$) 8.90 (1H, s), 8.41 (1H, s), 8.38 (1H, s), 8.33 (1H, s), 7.72 (1H, s), 7.45 (2H, s), 7.36 (2H, s); 6.86 (1H, s); 4.41 (2H, t, *J* 4.0 Hz); 4.07 (3H, s); 4.03 (3H, s); 3.52 (2H, q, *J* 4.0 Hz); 3.23 (2H, q, *J* 4.0 Hz); 2.60 (2H, t, *J* 4.0 Hz); 2.36 (6H, s), 2.01 (5H, m); LRMS (ES^+) m/z (rel intensity) 570 ($[M+H]^+$, 100%); HRMS [$M+H^+$, $C_{25}H_{35}N_{11}O_5^+$] calcd: 570.2909. Found: 570.2901.

f-IP(C_3gly)I (6)

A solution of f-IP(C_3 *N*-Cbz glycine)I (10 mg, 0.0139) was reduced in the presence of hydrogen over 10% Pd/C (10 mg) in cold MeOH (10 mL) for 2 days. The solution was filtered

through celite and the solvent removed by evaporation to yield glycine f-IPI (**6**) as a white solid (3 mg, 37%); R_f: [MeOH:$CHCl_3$:liq NH_3 (5:15:1)] 0.27; mp. 180 ^oC; IR: υ (ATR) cm^{-1} 2961, 1649, 1533, 1465, 1437, 1401, 1260, 1985, 1019, 894, 865, 767, 759, 693, 677, 666, 655, 623; NMR δ_H (400 MHz, CD_3OD) 8.25 (1H, s); 7.44 (2H, s); 7.02 (1H, s), 4.43 (2H, t, *J* 6Hz), 4.05 (3H, s), 4.02 (3H, s), 3.55 (4H, t, *J* 6Hz), 3.26 (2H, t, *J* 8 Hz), 2.80 (2H, t, *J* 6 Hz), 2.51 (6H, s), 2.03 (2H, quint, *J* 6 Hz); LRMS (ES^+) m/z (rel intensity) 585 ($[M+H]^+$, 25%), 293 (100%); HRMS [$M+H^+$, $C_{25}H_{36}N_{12}O_5^+$] calculated: 585.3008. Found: 585.3010.

5.4 Results

Synthesis of the target polyamides **1**-**6** were accomplished using solution phase aromatic amine – acid chloride or carboxylic acid coupling reactions. The synthesis of diamino f-ImPyIm (**1**) has recently been reported by us, and that involved catalytic hydrogenation of the corresponding azide f-ImPyIm polyamide (**3**)(*22*). Herein we report chloro f-ImPyIm (**2**) could be directly converted to diamino (**1**) through direct displacement of the chloride atom by ammonia via the S_N2 reaction. The isolated yield of (**1**) was 31%. The synthesis and characterization of polyamides **2** and **3** are given in the experimental section. The pyrrole-*N*1-alkylamino group in polyamide **1** was readily acylated by acetyl chloride to give acetamido f-ImPyIm (**5**) in 79% yield. Also, reaction of the amino group with *N*-Cbz-glycine in the presence of EDCI gave the protected glycine intermediate in 42% yield. Final removal of the Cbz group by catalytic hydrogenation afforded the desired glycine f-IPI polyamide (**6**) in 37% yield. The preparation of alkyne f-ImPyIm (**4**) required a 4-nitropyrrole-*N*1-pentyne synthon, which was coupled to a C-terminus imidazole moiety. Reduction of the nitro group using acidic stannous chloride and subsequent coupling of the amine with *N*-foramido-1-methylimidazole-2-carboxylic acid using PyBOP and DiPEA afforded the desired polyamide (4) in 24% yield.

Surface Plasmon Resonance

Surface plasmon resonance (SPR) biosensor experiments were performed to obtain quantitative measurements of the binding affinities, stoichiometries, and selectivities of polyamides **1-6**. Sensorgrams for the binding of polyamide **6** to the cognate sequence (5'ACGCGT-3') are shown as an example in Figure 5.2. The sensorgrams and fitting of the data show a strong, highly positive cooperative process for binding polyamide **6** to the sequence 5'-ACGCGT-3'with a concentration dependent observed on-rate, as expected, for a polyamide-DNA complexation process and a first-order, extremely slow off-rate. Steady-state RU values were obtained by equilibrium fitting and the observed RU_{max} values clearly indicate a 2:1 complex formation. The calculated binding constants for polyamide **6** were $K_1 = 2.4 \times 10^5$, $K_2 = 1.1 \times 10^9$, and $K_{eq} = 1.6 \times 10^7$, where $K_{eq} = \sqrt{K_1 \times K_2}$, and were determined using a 2:1 binding model. The binding constants for compounds **1-6** are listed in Table 5.1 and were calculated using the same 2:1 binding model. Polyamide **1** exhibited the strongest overall binding to the cognate DNA sequence, 5'-ACGCGT-3'. Binding affinities for polyamide **4** were not determined by SPR because the compound aggregated and exhibited surface problems with the biosensor flow cells, something occasionally observed with non-polar polyamides. The binding affinity and kinetic behaviors for polyamides **1-3**, **5**, and **6** are consistent for strong DNA binders. The rate of dissociation, k_d, for polyamide **6** is extremely slow as shown in Figure 5.2. Because of the extremely slow off-rate for polyamide **6**, an accurate value for k_d could not be determined by a kinetic fit of the data. This is also true for polyamides which have an extremely fast on-rate (k_a). At higher concentrations the rate of association (k_a) of polyamide **6** with its cognate DNA sequence (5'-ACGCGT-3') is too fast and a kinetic fit of the data would not produce an accurate value for k_a. However, it is possible to estimate an apparent half-life for the

dissociation of the polyamide-DNA complex from the sensorgrams. These are very useful for comparison and are included in Table 5.3. The apparent half-life for the dissociation of the dicationic polyamides, **1** and **6,** from the polyamide-DNA complex are >8.3 and 4.2 hours, respectively. The higher binding affinities for polyamides **1** and **6** are the direct result of the extremely slow rate of dissociation of the polyamides from the DNA complex.

Isothermal Titration Calorimetry

To better understand the interactions of polyamides **1-6** with the DNA sequence 5'-ACGCGT-3', ITC experiments were conducted and complete thermodynamic profiles were obtained (Table 5.1). An ITC titration thermogram and the corresponding integrated heats of the reaction for polyamide **6** with the sequence 5'ACGCGT-3' are shown as an example in Figure 5.3. The excess DNA method for ITC is employed to determine ΔH and not K_{eq} because some of the binding affinities are generally too high to be determined by ITC. Since binding affinities are determined via SPR, there is not a need to use ITC to determine that particular thermodynamic parameter. The excess DNA method for ITC allows for a very accurate determination of ΔH since a consistent measurement of Q (heat) over a molar range of ~0.6:1 (moles of DNA: moles ligand) allows for the precise linear fit and extrapolation back to zero mol ratio and the ΔH of the reaction (*27, 29*). The enthalpy of the reaction for polyamide **6** (ΔH) is -4.5 kcal/mol from the linear fit of the integrated heats shown in Figure 5.3. The Gibbs-Free energy (ΔG) for the interactions of polyamides **1-6** with the cognate DNA sequence 5'-ACGCGT-3' are collected with ΔH values in Table 5.1. ΔG of the reaction for polyamide **6** with 5'-ACGCGT-3' is favorable and was calculated to be -9.8 kcal/mol. TΔS for the interaction of polyamide **6** is calculated to be 5.3 kcal/mol using the ΔG value determined by SPR and the ΔH value determined by ITC. The binding of polyamide **6** to its cognate sequence is thus driven by a

combination of both favorable enthalpy and entropy contributions, though a slightly larger entropy contribution is noted. Polyamides **1** and **5** have very similar enthalpy and entropy contributions but are all enthalpically driven (Table 5.1).

Thermal Denaturation

Thermal denaturation studies were used as an initial evaluation of the affinity and sequence specificity of polyamides **1-6** using the DNA sequences 5'-ACGCGT-3' and 5'-A_3T_3-3'. Melting temperatures (ΔT_m) values were determined from the difference in melting temperatures of the DNA-polyamide complex and duplex DNA alone. ΔT_m for polyamides **1-6** for both sequences are shown in Table 5.2. Polyamides **1** and **6**, with cationic substituents, showed very high ΔT_m values of 24 and 20 °C for the sequence 5'-ACGCGT-3', respectively. ΔT_m values for polyamides **2-4** ~~and 4~~ with 5'ACGCGT-3'were less remarkable. All six polyamides showed little or no change in T_m value with the sequence 5'A_3T_3-3', as shown in Table 5.2.

DNase I Footprinting

DNase I footprinting studies provide the sequence selectivity of each polyamide synthesized for this study. The autoradiogram given in Figure 5.4 shows that a footprint for polyamide **2** at the 5-'ACGCGT-3' sequence is clearly evident at a concentration of 1 μM-. No footprints other than for the cognate sequence appear for polyamide **2**. The audiogram for polyamide **1** shows a footprint emerging at 0.01 μM and a clear footprint by 0.05 μM for the sequence 5'-ACGCGT-3' (Table 5.2). This supports the SPR results showing that polyamide **1**, the diamino derivative, exhibits the highest binding affinity to the cognate sequence 5'-

ACGCGT-3'. Table 5.2 lists the concentrations at which a clear footprint was evident for the sequence 5'-ACGCGT-3' for all six polyamides.

Circular Dichroism

Further investigation of the binding properties of polyamides **1-6** were evaluated using circular dichroism (CD) studies and the results for polyamide **6** are shown in Figure 5.5. The titration of polyamide **6** into the cognate oligonucleotide 5'-ACGCGT-3' produced a strong, positive DNA- induced band at ~330 nm which is indicative of the compound binding to the minor groove of DNA (*5, 6, 8, 17, 18, 22, 23*). A weaker induced CD band was observed for the noncognate sequence 5'-A_3T_3-3'. Compounds **1-5** also showed a DNA-induced band at 330 nM, thus indicating all these polyamides also bind to the minor groove of DNA. Further evidence that the polyamides are binding to the DNA via a single mechanism is the presence of an isodichroic point for the overlaid spectra at ~310 nM (shown for polyamide **6** in Figure 5.5), most likely as a stacked homodimer. This holds true for compounds **1-5** also.

5.5 Discussion

Since discovery and isolation of netropsin by Finlay *et al*. over sixty years ago (*31*), the study of netropsin, distamycin and their complexation with AT-rich sequences of the minor groove of DNA has become well recognized and accepted (*3, 5, 32-35*). The discovery that distamycin could form a "stacked", antiparallel dimer in the minor groove of DNA of varying AT sequences opened the door and created a rationale for synthetic polyamide drug design based on DNA sequence (*3, 5, 32, 33, 36-42*). The systematic definition of Im and Py containing synthetic polyamides binding rules for sequence specific base pair recognition of the minor groove for Watson-Crick duplex DNA by Dervan and co-workers has paved the way for the

development of an unlimited library of synthetic polyamides which can be programmed to bind to any particular DNA sequence (*5, 43-45*). Several groups have developed a wide range of synthetic polyamide monomers which bind to the minor groove of DNA as "stacked" dimers. Polyamides have shown excellent in-vitro activity, however, the development and progress of polyamides as therapeutics has been hindered by their inadequate solubility and cellular uptake, even as several positive in-vivo results have been reported (*9, 46-57*).

An underutilized development area is modification of the *N*5-position of the pyrrole and imidazole ring systems. While a number of changes have been reported at that position and H-pin polyamides are linked through the same *N*5-position of the center heterocycle, more detailed studies with systematic modifications are needed. To accomplish this goal we added alkyl cationic groups to the *N*5-position and have now extended the systematic analysis with the analogs shown in Figure 5.1.

The strategy is to functionalize polyamides through the addition of six different alkyl side chains at the *N*5-position of the central pyrrole heterocycle and monitor their impact on sequence selectivity and binding affinities. The importance of using f-ImPy*Im derivatives to target the sequence 5'-ACGCGT-3' is the presence of that particular sequence in the promoter region of the core sequence of the Mlu1 cell-cycle box which is a transcriptional element of the human Dbf4 gene. Again, the Dbf4 gene is a regulatory subunit of the Cdc7 kinase (a cyclin dependent kinase) which in high levels has been implicated in various cancers.

The parent polyamide, f-ImPyIm, has a very high binding affinity and of the analogs has a better affinity for 5'-ACGCGT-3' with the exception of polyamide **1** which is within experimental error of the parent compound (Table 5.1). Polyamides **1** and **6** demonstrated the

highest binding affinity of the new f-ImPyIm analogue compounds to the cognate sequence 5'-ACGCGT-3'with good sequence selectivity. Both of these polyamides have a second positive charge that contributes to the higher binding affinities compared to the other four polyamide derivatives. This is presumably due to an additional electrostatic interaction between the positively charged NH_3^+ present on the distal end of the alkyl side chain on both polyamides **1** and **6** and the negatively charged phosphate backbone of DNA. The length of the alkyl side chain allows it to extend out of the minor groove of the DNA and interact electrostatically with the phosphate back bone. The larger enthalpic contribution to the overall Gibbs-free energy of the reaction for polyamide **1** binding to the sequence 5'-ACGCGT-3' would support this explanation for the increased binding affinity. The cationic modification of polyamide **6** is more bulky than the substituent of polyamide **1** and this clearly causes some steric hindrance and reduced binding affinity. The uncharged compounds bind significantly more weakly than the parent f-ImPyIm as well as the dicationic compounds. This again probably represents some decrease in DNA interactions due to steric effects without the favorable interactions of a second cation.

Modulation of the rate of association and/or dissociation is another tool to modify and change the biological activity of polyamides. This study demonstrates that by making systematic modifications to polyamides, such as f-ImPyIm, the binding affinity can be altered through the modulation of k_d and/or k_a. More detailed work must be conducted to see if there are possible trends in the effects of modifying polyamides designed to target specific DNA sequences has on the rate of polyamide-DNA complex formation and complex dissociation.

The dicationic polyamides **1** and **6** demonstrated comparable binding affinities to that of the parent compound f-ImPyIm while having increased water solubility. Increased water

solubility addresses one of three major advancements that must be made in order for polyamides to be become potential therapeutics. While the dicationic compounds showed an increase in water solubility, the other two major areas of concern, sequence specificity and binding affinity, remained within experimental error of the parent compound, f-ImPyIm. The binding affinity for both dications remained comparatively the same to the parent compound because of the modulation of the rate of dissociation of the polyamide-DNA complex. Polyamides have the potential to control and modulate gene expression but in order to do so, they must readily enter the cells and localize in the nucleus. Along with increased water solubility, the two dicationic derivatives have more potential as therapeutic candidates over hairpin, H-pin and cyclic polyamides because of their smaller molecular weights and molecular design, thus, allowing them to penetrate into cells and localize in the nucleus.

In summary, there is a need to develop a library of modified polyamides which target DNA and exhibit improved biological properties. The present study examined the binding properties of six polyamide compounds, each with an attached alkyl side chain, with differing functional groups extending from the *N*5 position of the center heterocycle to the minor groove of the DNA sequence 5'-ACGCGT-3'.

5.6 Acknowledgements

The authors thank the NSF [CHE 0809162 and CHE 0922623 (NMR)], Cancer Research UK (C2259/A9994 to J.A.H.), and the Georgia Research Alliance for their generous support.

5.7 References

1. Mitra, S. N., Wahl, M. C., and Sundaralingam, M. (1999) Structure of the side-by-side binding of distamycin to d(GTATATAC)2, *Acta Crystallogr. Sect. D* 602-609.
2. Pelton, J. G., and Wemmer, D. E. (1990) Binding modes of distamycin A with d(CGCAAATTTGCG)2 determined by two-dimensional NMR, *J. Am. Chem. Soc. 112*, 1393-1399.
3. Pelton, J. G., and Wemmer, D. E. (1989) Structural characterization of a 2:1 distamycin A-d(CGCAAATTGGC) complex by two-dimensional NMR, *Proc. Natl Acad. Sci. 15*, 5723-5727.
4. O'Hare, C. C., Mack, D., Tandon, M., Sharma, S. K., Lown, J. W., Kopka, M. L., Dickerson, R. E., and Hartley, J. A. (2002) DNA sequence recognition in the minor groove by crosslinked polyamides: The effect of *N*-terminal head group and linker length on binding affinity and specificity, *Proc Natl Acad Sci U S A 99*, 72-77.
5. Lacy, E. R., Le, N. M., Price, C. A., Lee, M., and Wilson, W. D. (2002) Influence of a terminal formamido group on the sequence recognition of DNA by polyamides, *J. Am. Chem. Soc. 124*, 2153-2163.
6. Westrate, L., Mackay, H., Brown, T., Nguyen, B., Kluza, J., Wilson, W. D., Lee, M., and Hartley, J. A. (2009) Effects of the *N*-terminal acylamido group of imidazole and pyrrole-containing polyamides on DNA sequence specificity and binding affinity *Biochemistry 48*, 5679-5688.
7. O'Hare, C., Uthe, P., Mackay, H., Blackmon, K., Jones, J., Brown, T., Nguyen, B., Wilson, W. D., Lee, M., and Hartley, J. A. (2007) Sequence recognition in the minor groove of DNA by covalently linked formamido Imidazole-Pyrrole-Imidazole polyamides: Effect of H-Pin linkage and linker length on selectivity an d affinity, *Biochemistry 46*, 11661-11670.
8. Buchmueller, K. L., Bailey, S. L., Matthews, D. A., Taherbhai, Z. T., Register, J. K., Davis, Z. S., Bruce, C. D., O'Hare, C., Hartley, J. A., and Lee, M. (2006) Physical and structural basis for the strong Interactions of the -ImPy- central pairing motif in the polyamide f-ImPyIm, *Biochemistry 45*, 13551-13565.
9. Chou, C. J., Farkas, M. E., Tsai, S. M., Alvarez, D., Dervan, P. B., and Gottesfeld, J. M. (2008) Small molecules targeting histone H4 as potential therapeutics for chronic myelogenous leukemia, *Mol. Cancer Ther 7*, 769-778.

10. Geierstanger, B. H., Mrksich, M., Dervan, P. B., and Wemmer, D. E. (1996) Extending the recognition site of designed minor groove binding molecules., *Nat. Struct. Biol. 3*, 321-324.
11. Chenoweth, D. M., and Dervan, P. B. (2009) Allosteric modulation of DNA by small molecules., *Proc. Natl. Acad. Sci. 106*, 13175-13179.
12. Hochhauser, D., Kotecha, M., O'Hare, C., Morris, P. J., Hartley, J. M., Taherbhai, Z., Harris, D., Forni, C., Mantovani, R., and Lee, M. H., J. A. . (2007) Modulation of topoisomerase IIalpha expression by a DNA sequence-specific polyamide, *Mol. Cancer Ther. 6*, 346-354.
13. Dwyer, T. J., Geierstanger, B. H., Bathini, Y., Lown, J. W., and Wemmer, D. E. (1992) Design and binding of a distamycin A analog to d(C G C A A G T T GG C) • d(G C C A A C T T G C G): Synthesis, NMR studies, and implications for the design of sequence-specific minor groove binding oligopeptides, *J. Am. Chem. Soc. 114*, 5911-5919.
14. Shinohara, K., Bando, T., and Sugiyama, H. (2010) Anticancer activities of alkylating pyrrole-imidazole polyamides with specific sequence recognition., *Anticancer Drugs 21*, 228-242.
15. Minoshima, M., Bando, T., Shinohara, K., and Sugiyama, H. (2009) Molecular design of sequence specific DNA alkylating agents., *Nucleic Acids Symp. Ser. 23*, 69-70.
16. Borman, S. (2010) New route to amide formation, *Chem. Eng. News 88*, 13.
17. Buchmueller, K. L., Staples, A. M., Howard, C. M., Horick, S. M., Uthe, P. B., Le, N. M., Cox, K. K., Nguyen, B., Pacheco, K. A., Wilson, W. D., and Lee, M. (2005) Extending the language of DNA molecular recognition by polyamides: unexpected influence of imidazole and pyrrole arrangement on binding affinity and specificity, *J. Am. Chem. Soc. 127*, 742-750.
18. Buchmueller, K. L., Staples, A. M., Uthe, P. B., Howard, C. M., Pacheco, K. A., Cox, K. K., Henry, J. A., Bailey, S. L., Horick, S. M., Nguyen, B., Wilson, W. D., and Lee, M. (2005) Molecular recognition of DNA base pairs by the formamido/pyrrole and formamido/imidazole pairings in stacked polyamides, *Nucleic Acids Res. 33*, 912-921.
19. Verma, R., Patapoutian, A., Gordon, C. B., and Campbell, J. L. (1991) Identification and purification of a factor that binds to the Mlu I cell cycle box of yeast DNA replication genes, *Proc. Natl. Acad. Sci. 88*, 7155-7159.
20. Wu, X., and Lee, H. (2002) Human Dbf4/ASK promoter is activated through the Sp1 and MluI cell-cycle box (MCB) transcription elements., *Oncogene 21*, 7786-7796.
21. Yamada, M., Sato, N., Taniyama, C., Ohtani, K., Drong, R. F., Weiland, K. L., Slightom, J. L., Scalafani, R. A., and Hollingsworth, R. E. (1998) A human homolog of the yeast CDC7 gene is overexpressed in some tumors and transformed cell lines, *Gene 211*, 133-140.
22. Babu, B., Liu, Y., Plaunt, A., Riddering, C., Ogilvie, R., Westrate, L., Davis, R., Ferguson, A., Mckay, H., Rice, T., Ramos, J. P., Chavda, S., Wilson, W. D., Lin, S., Kiakos, K., Hartley, J. A., and Lee, M. (2011) Design, synthesis and DNA binding properties of orthogonally positioned diamino containing polyamide f-IPI., *Biochem Biophys Res Commun. 404*, 848-852.
23. Satam, V., Babu, B., Chavda, S., Savagian, M., Sjoholm, R., Tzou, S., Ramos, J. P., Liu, Y., Kiakos, K., Lin, S., Wilson, D. W., Hartley, J. A., and Lee, M. (2012) Novel diamino imidazole and pyrrole-containing polyamides: Synthesis and DNA binding studies of

mono- and diamino-phenyl-ImPy*Im polyamides designed to target 5'-ACGCGT-3', *Bioorg. Med. Chem. 22*, 693-701.
24. Nguyen, B., Tanious, F., and Wilson, W. D. (2006) Biosensor-surface plasmon resonance: Quantitative analysis of small molecule–nucleic acid interactions, *Methods 42*, 150-161.
25. Lacy, E. R., Cox, K. K., Wilson, W. D., and Lee, M. (2002) Recognition of T*G mismatched base pairs in DNA by stacked imidazole-containing polyamides: Surface plasmon resonance and circular dichroism studies, *Nucleic Acids Res. 30*, 1834-1841.
26. Wanunu, M., and Tor, Y., (Eds.) (2012) *Methods for Studying Nucleic Acid/Drug Interactions*, CRC Press, Boca Raton, FL.
27. Haq, I., Jenkins, T. C., Chowdhry, B. Z., Ren, J., and Chaires, J. B. (2000) Parsing free energies of drug-DNA interactions, *Methods Enzymol 323*, 373-405.
28. Chaires, J. B. (2006) A thermodynamic signature for drug-DNA binding mode, *Arch Biochem Biophys 453*, 26-31.
29. Ren, J., Jenkins, T. C., and Chaires, J. B. (2000) Energetics of DNA intercalation reactions, *Biochemistry 39*, 8439-8447.
30. Mulder, K., Sexton, J., Taherbhai, Z., Jones, J., Uthe, P., Brown, T., and Lee, M. (2008) N-formamido-containing mono- and diheterocyclic pyrrole-and imidazole-2-carboxylic acids as building blocks for polyamide synthesis, *Synth. Commun. 38*, 33-44.
31. Finlay, A. C., Hochstein, F. A., Sobina, B. A., and Murphy, F. X. (1951) Netropsin, a new antibiotic produced by a Streptomyces, *J. Am. Chem. Soc. 73*, 341-343.
32. Wemmer, D. E. (2001) Ligands recognizing the minor groove of DNA: Development and applications, *Biopolymers 52*, 197-211.
33. Wemmer, D. E. (2000) Designed sequence-specific minor groove ligands, *Annu. Rev. Biophys. Biomol. Struct. 29*, 439-461.
34. Bailly, C., and Chaires, J. B. (1998) Sequence-specific DNA minor groove binders. Design and synthesis of netropsin and distamycin analogues, *Bioconjug. Chem. 9*, 513-538.
35. Neidle, S. (1997) Crystallographic insights into DNA minor groove recognition by drugs, *Biopolymers 44*, 105-121.
36. Wemmer, D. E., Geierstanger, B. H., Fagan, P. A., Dwyer, T. J., P., J. J., Pelton, J. G., Ball, G. E., Leheny, A. R., Chang, W.-H., Bathini, Y., Lown, J. W., Rentzeperis, D., Marky, L. A., Singh, S., and Kollman, P., (Eds.) (1993) *Minor groove recognition of DNA by Distamycin and its analogs*.
37. Bailly, C., (Ed.) (1998) *Advances in DNA Sequence Specific Agents*, Vol. 3, JAI Press Inc., Greenwich, CT.
38. Lown, J. W. (1993) Targeting the DNA Minor Groove for Control of Biologiocal Function: Progress, Challenges and Prospects., *Chemtracts: Org. Chem 6*, 205-237.
39. Urbach, A. R., and Dervan, P. B. (2001) Toward rules for 1:1 polyamide:DNA recognition, *Proc. Natl. Acad. Sci. U S A 98*, 4343-4348.
40. Lee, M., Rhodes, A. L., Wyatt, M. D., Forrow, S., and Hartley, J. A. (1993) GC base sequence recognition by oligo(imidazolecarboxamide) and C-terminus-modified analogues of distamycin deduced from circular dichroism, proton nuclear magnetic resonance, and methidiumpropylethylenediaminetetraacetate-iron(II) footprinting studies, *Biochemistry 32*, 4237-4245.

41. Yang, X. L., Kaenzig, C., Lee, M., and Wang, A. H. (1999) Binding of AR-1-144, a tri-imidazole DNA minor groove binder, to CCGG sequence analyzed by NMR spectroscopy, *Eur. J. Biochem. 263*, 646-655.
42. Yang, X. L., Hubbard, R. B., IV., Lee, M., Tao, Z. F., Sugiyama, H., and Wang, A. H. (1999) Imidazole-imidazole pair as a minor groove recognition motif for T:G mismatched base pairs, *Nucleic Acids Res. 27*, 4183-4190.
43. Dervan, P. B., and Burli, R. W. (1999) Sequence-specific DNA recognition by polyamides, *Curr. Opin. Chem. Biol. 3*, 688-693.
44. Wemmer, D. E., and Dervan, P. B. (1997) Targeting the minor groove of DNA, *Curr Opin Struct Biol 7*, 355-361.
45. White, S., Baird, E. E., and Dervan, P. B. (1997) On the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides, *Chem. Biol. 4*, 569-578.
46. Jacobs, C. S., and Dervan, P. B. (2009) Modifications at the C-terminus to improve pyrrole-imidazole polyamide activity in cell culture., *J. Med. Chem. 52*, 7380-7388.
47. Nickols, N., Jacobs, C., Farkas, M., and Dervan, P. (2007) Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface, *ACS Chem. Bio. 2*, 561-571.
48. Nickols, N., and Dervan, P. (2007) Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide, *Proc Natl Acad Sci 104*, 10418-10423.
49. Burnett, R., Melander, C., Puckett, J. W., Son, L. S., Wells, R. D., Dervan, P. B., and Gottesfeld, J. M. (2006) DNA sequence-specific polyamides alleviate transcription inhibition associated with long GAA.TTC repeats in Friedreich's ataxia, *Proc. Natl. Acad. Sci. 103*, 11497-11502.
50. Puckett, J. W., Muzikar, K. A., Tietjen, J., Warren, C. L., Ansari, A. Z., and Dervan, P. B. (2007) Quantitative microarray profiling of DNA-binding molecules, *J. Am. Chem. Soc. 129*, 12310-12319.
51. Nickols, N., Jacobs, C. S. F., M. E., and Dervan, P. B. (2007) Improved nuclear localization of DNA-binding polyamides, *Nucleic Acids Res. 35*, 363-370.
52. Edelson, B. S., Best, T. P., Olenyuk, B., Nickols, N. G., Doss, R. M., Foister, S., Heckel, A., and Dervan, P. B. (2004) Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates, *Nucleic Acids Res 32*, 2802-2818.
53. Hsu, C. F., and Dervan, P. B. (2008) Quantitating the concentration of Py-Im polyamide-fluorescein conjugates in live cells, *Bioorg. Med. Chem. Lett. 18*, 5851-5855.
54. Best, T. P., Edelson, B. S., Nickols, N. G., and Dervan, P. B. (2003) Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture, *Proc Natl Acad Sci U S A 100*, 12063-12068.
55. Harki, D. A., Satyamurthy, N., Stout, D. B., Phelps, M. E., and Dervan, P. B. U. S. A., 105, 13039. (2008) In vivo imaging of pyrrole-imidazole polyamides with positron emission tomography, *Proc. Natl. Acad. Sci. 105*, 13039-13044.
56. Fukasawa, A., Aoyama, T., Nagashima, T., Fukuda, N., Ueno, T., Sugiyama, H., Nagase, H., and Matsumoto, Y. (2009) Pharmacokinetics of pyrrole-imidazole polyamides after intravenous administration in rat, *Biopharm. Drug Dispos. 30*, 81-88.
57. Matsuda, H., Fukuda, N., Ueno, T., Tahira, Y., Ayame, H., Zhang, W., Bando, T., Sugiyama, H., Saito, S., Matsumoto, K., Mugishima, H., and Serie, K. (2006) Development of gene silencing pyrrole-imidazole polyamide targeting the TGF-beta1 promoter for treatment of progressive renal diseases, *J. Am. Soc. Nephrol. 17*, 422-432.

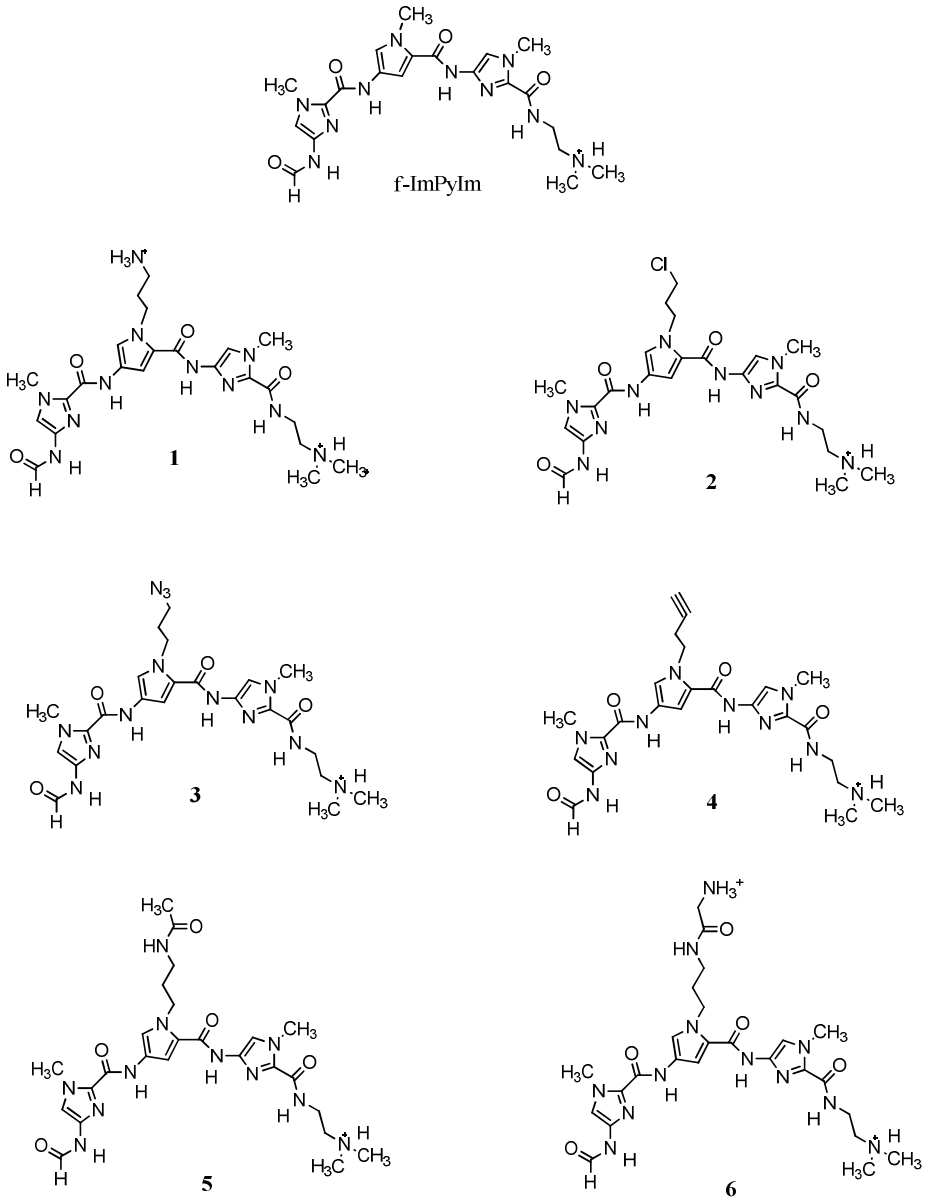


Figure 5.1 Structures of the "parent" compound (f-ImPyIm), f-ImPy($(CH_2)_3NH_3$)Im (**1**), f-ImPy($(CH_2)_3Cl$)Im (**2**), f-ImPy($(CH_2)_3N_3$)Im (**3**), f-ImPy($(CH_2)_3$alkyne)Im (**4**), f-ImPy($(CH_2)_3$NHAc)Im (**5**), and f-ImPy($(CH_2)_3$NHGly)Im (**6**).

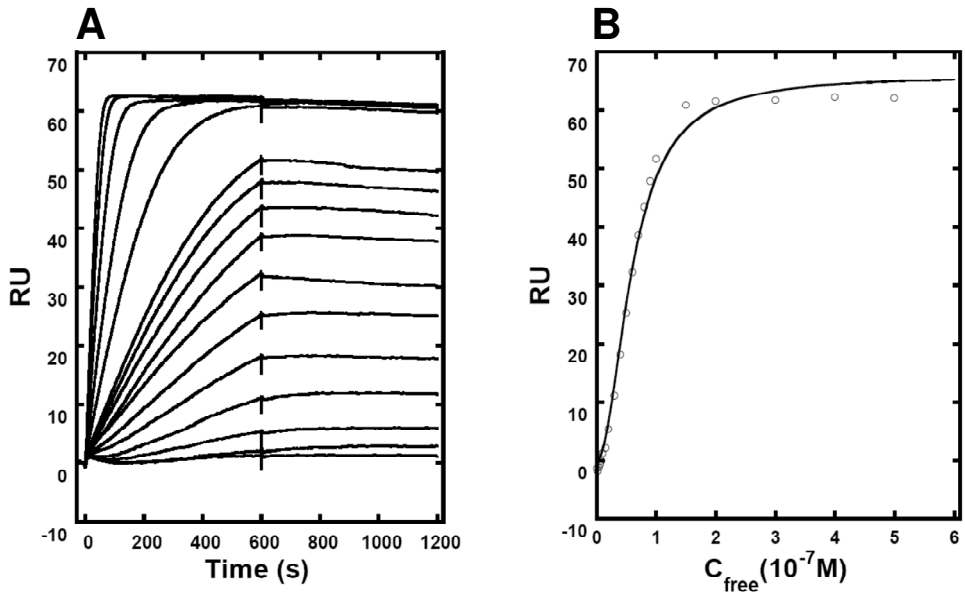


Figure 5.2 SPR sensorgrams for the binding of f-ImPy($(CH_2)_3$NHGly)Im (**6**) to the sequence 5'-ACGCGT-3'**(A)** and a plot of RU_{ss} *vs* concentration including a best-fit line **(B)**. The cooperative ($K_2 > K_1$) binding of two f-ImPy($(CH_2)_3$NHGly)Im molecules per 5'-ACGCGT-3' is observed.

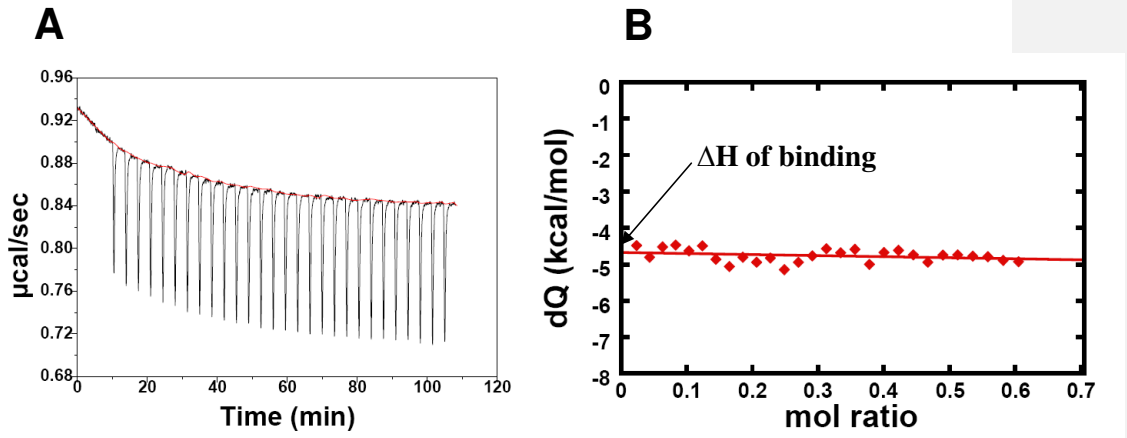
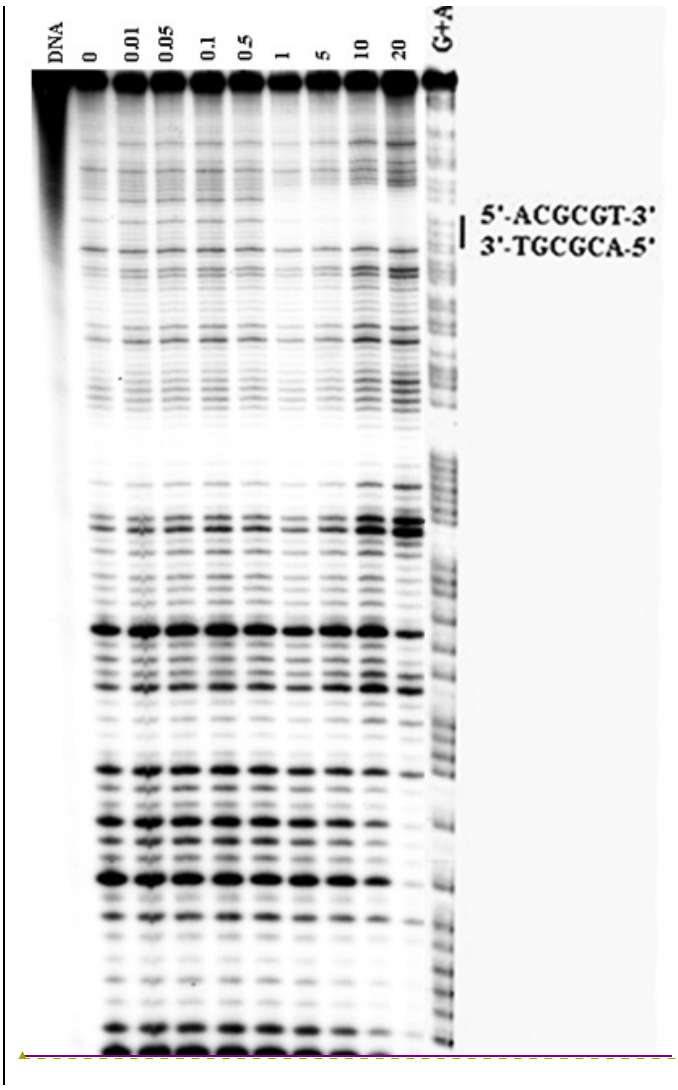


Figure 5.3 Experimental ITC data **(A)** for the excess DNA method titration of polyamide **6** into 5'-ACGCGT-3' and the integrated heats for each injection of **6** used for the titration (red dots) and a linear fit (red line) of that data which is extrapolated to zero and yields of the ΔH reaction **(B)**.



Formatted: Font: (Default) Times New Roman, 12 pt

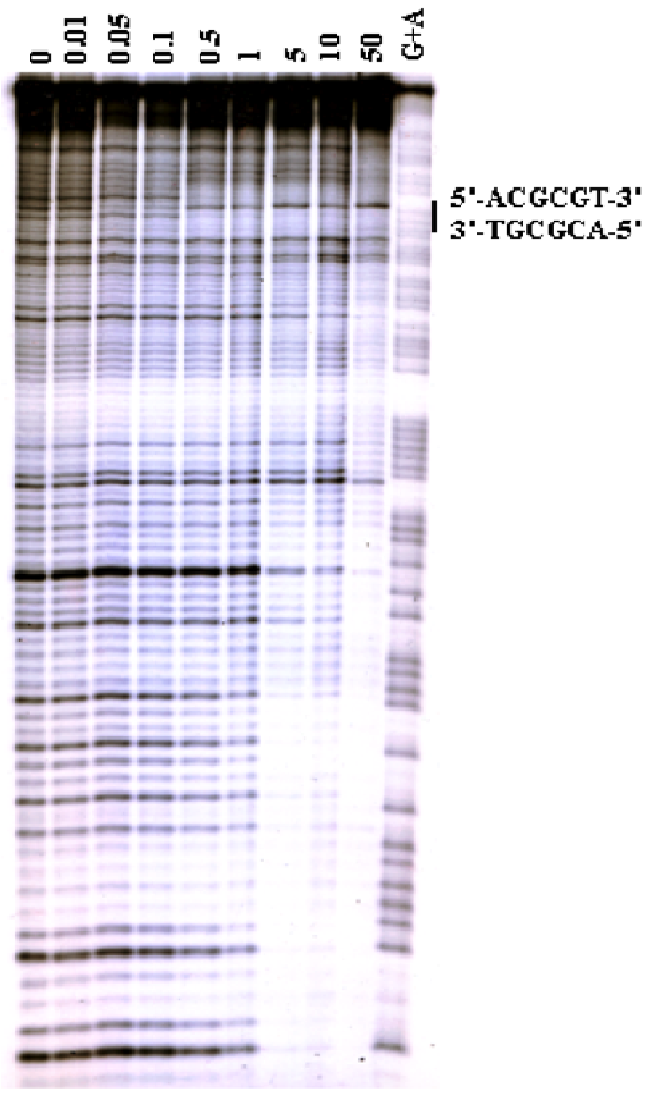


Figure 5.4 DNase I footprinting of compound **2** (f-ImPy($(CH_2)_3$CI)Im) on the antisense strand of the 5'-[^{32}P] 130-bp fragment. All reactions contain ~500cps DNA fragment, 10 mM Tris, pH 7, 1 mM EDTA, 50 mM KCL, 1 mM $MgCl_2$, 0.5 mM DTT, 20 mM HEPES. G + A denotes the purine sequencing lane. The 5'-ACGCGT-3' site is indicated by a solid bar.

Formatted: Font: (Default) Times New Roman, 12 pt

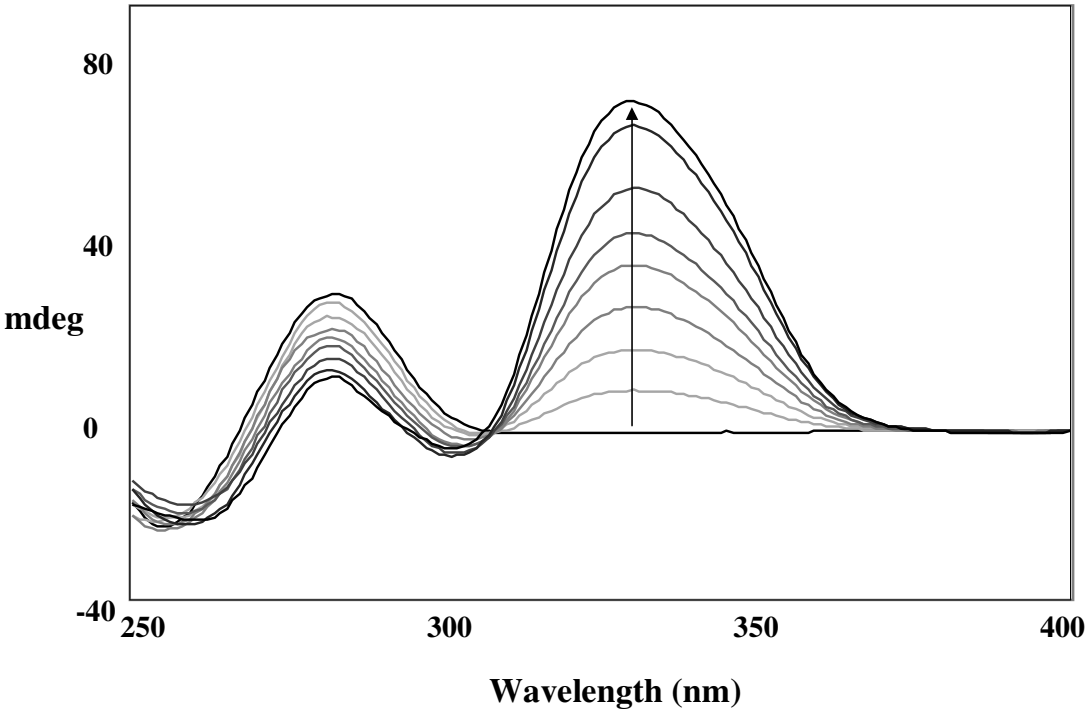


Figure 5.5 CD data for compound **6**, (f-ImPy($(CH_2)_3$NHGly)Im), with the sequence 5'-ACGCGT-3'. CD experiments were carried out using 160 μL of a 9 μM DNA solution, which was titrated with 1 mol equivalents of **6,** past the point of saturation. The spectra represent titration from DNA alone to 1 mol equivalent of polyamide **6** to the DNA hairpin, 2, 3, 4, 5, 6, 8 and 10 molar equivalents.

Table 5.1 Results for ITC and SPR experiments on polyamides **1-3, 5 and 6** with the sequence 5'-ACGCGT-3'. Included at the bottom are the thermodynamic properties and K_{eq} value for the "parent" compound, f-ImPyIm, as where previously determined and from which these modified polyamides were derived.

Polyamide	**ΔH[a] (kcal/mol)**	**ΔG[b] (kcal/mol)**	**TΔS[c] (kcal/mol)**	**$K_{eq}(M^{-1})$[b]**
1	-5.9	-11.4	5.5	2.8×10^8
2	-4.4	-8.3	3.9	1.2×10^6
3	-4.5	-9.8	5.3	1.5×10^6
5	-5.7	-9.1	3.4	4.9×10^6
6	-4.5	-10.6	6.1	6.2×10^7
f-ImPyIm(*5*)	-7.6	-11.3	3.7	1.9×10^8

$K_{eq} = \sqrt{\mathbf{K_1} \times \mathbf{K_2}}$; UD = unable to be determined due to insufficient heats from ITC experiments. [a]Determined by ITC. [b]Determined from SPR studies. [c] Calculated from the equation ΔS = (ΔG-ΔH)/T, using ΔG calculated from SPR and ΔH calculated from ITC. K values for polyamide 4 were unable to be determined by SPR because of surface problems.

Table 5.2 Results for thermal denaturation and DNase I footprinting experiments for polyamides **1-6** and the parent compound f-ImPyIm from previously reported work.

	ΔT_m (°C)		
Polyamide	**5'-ACGCGT-3'**	**5'-A_3T_3-3'**	**Footprinting (μM)[a]**
1	>24	10	0.05
2	1.5	0	1
3	4.8	0	1
4	0	0	10
5	10	1	0.1
6	20	5	0.5
f-ImPyIm (*22*)	7.3	NA	0.05

NA- not available. [a]Concentration when footprint at the cognate site became apparent.

Table 5.3 Estimated half-lives (in hours) for the rate of dissociation of polyamides 1-3, 5 and 6 from the sequence 5'-ACGCGT-3'.

Polyamide	**~$T_{1/2}$ (app*)** (hr)
1	>8.3
2	0.22
3	0.57
5	2.6
6	4.2

* apparent half-lives based on dissociation rates from SPR experiments.

CHAPTER 6:

PROMOTER SCANNING OF THE HUMAN COX-2 GENE WITH 8-RING POLYAMIDES: UNEXPECTED WEAKENING OF POLYAMIDE-DNA BINDING AND SELECTIVITY BY REPLACING AN INTERNAL N-Me-PYRROLE WITH β-ALANINE

The work presented in this chapter is based on a submitted paper entitled "Promoter scanning of the COX-2 gene with 8-ring polyamides: Unexpected weakening of the polyamide-DNA binding and selectivity by replacing an internal N-Me-pyrrole with β-alanine". This paper has been submitted to *Biochimie* as of April of 2012. Synthetic, fluorescence, and DNase footprinting methods, along with the corresponding written portions of this manuscript are ascribed to Dr. James K. Bashkin of the University of Missouri, St. Louis and his research group. My contribution to this paper included the SPR experiments, data analyses and writings, with the combined efforts of Shuo Wang and Drs. Rupesh Nanjunda and W. David Wilson.

Promoter Scanning of the Human COX-2 Gene with 8-Ring Polyamides: unexpected weakening of polyamide-DNA Binding and selectivity by replacing and internal N-Me-Pyrrole with a *β*-alanine

James K. Bashkin[†], Karl Aston[†], Joseph P. Ramos[‡], Kevin J. Koeller[†], Shuo Wang[‡], Rupesh Nanjunda[‡], Gaofei He[†], Cynthia M. Dupureur[†], and W. David Wilson[‡]

[†]*Department of Chemistry & Biochemistry, Center for Nanoscience, University of Missouri-St. Louis, One University Blvd., St. Louis, MO 63121, USA*

[‡]*Department of Chemistry, Georgia State University, Atlanta, GA 30302, USA*

6.1 Abstract

Rules for polyamide DNA recognition have proved invaluable for the design of sequence-selective DNA-binding agents in cell-free systems. However, these rules are not fully transferrable to predicting activity in cells, tissues or animals, and some refinements in understanding DNA recognition would help biomedical studies. Similar complexities are encountered when using internal β-alanines as polyamide building blocks in place of N-methyl pyrrole; β-alanines were introduced in polyamide designs to maintain good hydrogen bonding registry with the target DNA, especially for long polyamides (P.B. Dervan, A.R. Urbach, Essays Contemp. Chem. (2001) 327-339). Thus, to expand current rules, we compare the effects of replacing a single pyrrole with β-alanine in 8-ring polyamides. Replacement of a single internal N-methylpyrrole with β-alanine in two 8-ring polyamides causes a decrease in DNA binding affinity by two orders of magnitude and decreases DNA binding selectivity, contrary to expectations. Measurements were made by fluorescence spectroscopy, quantitative DNA footprinting and surface plasmon resonance, with these vastly different techniques showing excellent agreement. Furthermore, results were validated for a range of DNA substrates from small hairpins to long dsDNA sequences. These results combine with binding rules to refine design principles and expectations for polyamide-DNA recognition.

Keywords: Polyamide, DNA, fluorescence spectroscopy, surface plasmon resonance, footprinting, capillary electrophoresis.

6.2 Introduction

N-methylpyrrole and -imidazole based polyamides (PA) are being used with increasing frequency in fundamental and applied biomedical research programs [1-4]. Although the field owes much to the work on lexitropsins and their analogs [5], most of the recent improvements in polyamide utility derive from the DNA recognition rules developed by Dervan [6]. These rules in principle allow one to build a polyamide that recognizes a DNA sequence of interest, for example in order to disrupt transcription factor-DNA binding as a means of controlling gene expression [7]. Dervan has also published extensively on refinements to the understanding of preferred and less-favored polyamide-DNA interactions [8-10].

Even with this extensive knowledge at hand and the important additions to the field by Sugiyama, Laemmli, Lee and others [11-17], there is much that is unknown about PA design and modes of PA action in living cells. For example, we have observed large variations in antiviral activity for polyamides that ostensibly bind the same DNA sites and are in some cases isomers of each other [18]. This work introduced a new approach to the use of polyamides: we did not try to disrupt transcription factor binding, but rather designed polyamides to bind protein/ DNA recognition sites such as those important for viral replication and maintenance of the double-stranded viral DNA genome of high-risk, or cancer causing, human papillomavirus (HPV). Furthermore, we have found some complications in employing the process called promoter scanning [19]: the method can identify hotspots for DNA-PA interactions that lead to improved control of gene expression. If these hotspots are near each other, it is tempting to construct larger PA molecules from the smaller species identified in the promoter scanning process in order to take advantage of improved DNA binding strength and selectivity. This approach is particularly relevant in light of Sugiyama's report that uptake of polyamides into some

mammalian cells can be effective and independent of MW over the range from 400 – 4,000 [11]. Such results agree with our own observations that PAs with MWs well above 2000 show significant anti-HPV activity in human keratinocyte monolayers and in organotypic tissue culture [18]. However, we have not seen good correspondence between the activity of shorter polyamides and longer species prepared by combining short, active species. One complication common to these matters is the question of how best to introduce β-alanine (β) groups [20] to act as the molecular springs necessary to keep good registry between the hydrogen bonding groups of DNA and polyamides as the length of polyamides increases.

In order to better understand the incorporation of β in the construction of functional PAs from smaller molecules, we have carried out a systematic study of the role of β/N-methylpyrrole (Py) pairs vs. Py/Py pairs for DNA base pair recognition and binding. The literature provides systematic studies of β/β pairs vs. Py/Py pairs [20, 21], but not as much information is available on any changes to DNA binding that might occur by replacing a single Py residue with β. We report that, in contrast to the manner in which β helps maintain tight and selective DNA binding for some PA designs, especially when used as β/β pairs, β/Py pairs can greatly decrease the binding affinity and specificity of PA molecules for DNA.

In order to validate our findings, we used orthogonal assay methods for measuring DNA binding thermodynamics and kinetics, including a fluorescence assay [22], Biacore using biotin-labeled DNA hairpins [23-25], and quantitative footprinting [26]. We also studied a variety of DNA substrates from hairpins with 10 or 20 bp of duplex DNA to a linear, 120-mer dsDNA with one PA binding site and a 524-mer dsDNA with multiple PA binding sites. Quantitative footprinting was done using fluorescent labels and capillary electrophoresis (CE) [27] rather than more traditional radiolabeling and PAGE methods.

6.3 Materials and Methods

6.3.1 Synthesis

Polyamides were prepared by solid phase synthesis using Boc methodology [28] and were characterized by HPLC/MS (ESI+), ^{1}H NMR, high resolution mass spectrometry (ESI$^+$ MS^n) and combustion analysis. Of note, isolation by RP-HPLC using mobile phases containing 0.1%TFA led to all basic nitrogens being protonated, including imidazole. This contrasts with a previous report that used similar isolation procedures but indicated that imidazole groups were not protonated. Elemental analysis for C, H, N and F has been vital to our characterization efforts and our determination of molecular formulae.

6.3.2 Fluorescence Spectroscopy

Changes in fluorescence intensity of a TAMRA-labeled, hairpin-forming oligonucleotide as a function of PA concentration were used to quantitate DNA binding [22]. 5'-CCT GGA GAG GAA GCC AAG TGT TTT CAC TTG GCT TCC TCT CCA GG-3' was purchased from IDT HPLC pure (Coralville, IA) either unlabeled or labeled with T-TAMRA at either T_{34} or T_{37} as noted by the underlined T positions in the above sequence and as shown by asterisks in Figure 6.2. Using a Centricon unit, the DNA was rinsed twice with Milli-Q water, subsequently annealed from boiling water, quantitated using the vendor extinction coefficient, and either aliquotted, lyophilized and stored at -20 °C or used directly.

The experiments were performed using quartz cuvettes at 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, pH 7.4, 25 °C on a Fluorolog-3 (SPEX) spectrofluorimeter. All samples

were stirred continuously. TAMRA-labeled oligonucleotides were excited at 559 nm and the resulting emission observed through a monochromator set at 580 nm or a 592 nm bandpass filter (Edmund Optics, Barrington, NJ). Intensity values were obtained in triplicate and averaged and then normalized to fraction of bound DNA and then plotted against PA concentration. Resulting data were fit to equation **Eq. 1**:

$$\Theta = \frac{K_a[L]}{1 + K_a[L]} \qquad \text{(Eq. 1)}$$

where Θ is fraction of bound duplex, [L] is the total concentration of PA, and K_a is the association constant. K_d values represent an average of at least three separate experiments.

6.3.3 Competition Fluorescence Binding Assay

Aliquots of unlabeled DNA hairpin PA were added to a solution of fluorescent DNA-PA complex at a concentration 10-fold above the K_d. The resulting data were normalized and fit to a mathematical model describing both the labeled PA-DNA equilibrium (known K_d) and the unlabeled PA-DNA equilibrium (unknown K_d) using Scientist software (MicroMath, St. Louis, MO) as previously described [22].

6.3.4 Surface Plasmon Resonance (SPR)

Biosensor-SPR experiments were conducted with a BIAcore T100 or T200 instrument using similar procedures to those previously described [23, 25]. 5'-Biotin labeled DNA hairpins were immobilized on streptavidin-linked sensor chips via non-covalent capture. The tight affinity

between biotin and streptavidin yields a highly stable surface over time, allowing for regeneration of DNA surfaces with relatively harsh conditions. Immobilization in each flow cell is performed independently in a separate cycle, so that different DNAs can be used in different flow cells. After washing the chip with 1 M NaCl/50 mM NaOH and buffer to remove unlinked streptavidin, a 50 nM solution of each DNA sequence was injected over a streptavidin derivatized flow cell until ~400 resonance units (RU) of the long DNA (complete) sequence (as shown in Figure 6.2) was immobilized and ~250 RU of shorter cognate and noncognate sequences were immobilized.

Three flow cells contained DNA and one flow cell was left blank as a reference. For binding studies PA1-4 were diluted to different concentrations in degassed and filtered HEPES buffer (0.01 M HEPES and 0.001 M EDTA, 0.05 M NaCl, pH 7.4, and 0.05% surfactant P-20) and the diluted samples were injected over the DNA surface for a selected time. A 10 mM glycine solution at pH 2.5 was used for flow cell surface regeneration. Kinetics fits and steady-state binding studies were carried out as previously described [23, 25]. Steady-state analysis was done by averaging the RU values (RU) in the plateau region of the sensorgrams over a selected time region at different compound concentrations. The binding constants were obtained from fitting plots of steady-state RU versus C_{free} while kinetics fits were done in both the association and dissociation time regions [23, 25].

6.3.5 Quantitative DNase I Footprinting by Capillary Electrophoresis

A 120 bp duplex was generated as described previously [22]. Its sequence is: Fam_ATGTGATAGGGTAGATGATGGAGGTGgagtttaatgaaatttctgcaagggtctgtaatatgttttgtaaattc taa CCTGGAGAGG**AAGCCA**AGTG tttttggttacaaccattagcag. The first sequence in capital

letters is from HPV16 2150-2175. The sequence in lower case letters originates from HPV16 positions 2327-2377 and 2384-2406, which are the sequences surrounding the 5'-AAGCCA-3' target site (bolded). The underlined sequence is the same as the stem region of the DNA hairpin used for the above fluorescence studies and most SPR studies (i.e. the "long hairpin"). To generate a 524 bp DNA fragment of the HPV16 genome containing three cognate sites for PA1 and beginning at nt 2150, two oligomers were purchased from Integrated DNA Technologies, Inc. (Coralville, Iowa 52241). The forward primer sequence is: 5'-Fam-ATG TGA TAG GGT AGA TGA TGG AGG TG; the reverse primer sequence is: 5'-G CTC ATA CAC TGG ATT TCC GTT TTC GTC ATA. pUC19 with an HPV16 genome insert (Accession # AF125673) was used as the template. PCR settings were 55 ºC/1 min annealing temperature, 68 ºC/min extension temperature/time with 33 cycles. The fragment was purified using a 2% agarose gel and purified with Qiagen gel extraction kit (Qiagen, Valencia, CA 91355).

For footprinting [26], duplex DNA was mixed with polyamide in TKMC buffer (10 mM Tris, 10 mM KCl, 5 mM $MgCl_2$ and 5 mM $CaCl_2$), incubated at 37 ºC for time periods varying from 4 h to overnight. PA concentrations as high as 60-80 nM were used to study weak binding sites. Similar footprinting results were obtained for all these incubation times. The mixture was digested with RQ1 Rnase-Free DNase I (Promega, Madison, WI) for 5 min at 37 ºC and quenched by adding 30 µL of 500 mM EDTA. The DNase I digested product was purified with Qiagen PCR purification kit (Qiagen, Valencia, CA) and eluted with 30 µL of buffer. A few µL of the resulting samples were analyzed using an ABI 3100 Capillary Electrophoresis sequencer (Carlsbad, California). Data were processed using Genemarker V1.97 software (Softgenetics LLC, State College, PA). Peaks in the footprint were normalized to a neighboring peak not

sensitive to PA concentration, plotted as fraction bound vs. PA concentration, and fit to the Langmuir isotherm as described above.

6.4 Results

6.4.1 Polyamide and DNA Target Design

To carry out the comparative biophysical studies of the effects on DNA binding of β/Py replacements at internal positions of eight-ring polyamides, we prepared the compound pairs shown in Figure 6.1. The resulting compounds all target part of the human COX-2 Ets-1 binding site [29-31].

Compounds PA1-4 comprise two isomer pairs (PA1/PA2 and PA3/PA4) that differ only by replacement of an internal Py building block by β-alanine. All four compounds were designed to recognize the same sequence, but they bind in different orientations, as shown in Figure 6.2. The essentially identical DNA hairpins used in this study are based on the human COX-2 promoter in the Ets-1 binding region: the only difference is the location of a TAMRA dye for fluorescence studies, at T_{37} for PA1-2 and at T_{34} for PA3-4. These hairpins were used in the present study as (a) dye derivatives, (b) unfunctionalized DNA and (c) 5'-biotin conjugates in order to measure DNA-PA dissociation constants and kinetic parameters of PA1-4 binding their cognate DNA sites. In addition, the binding of PA1 and PA2 was studied by quantitative footprinting on a linear 120-bp DNA duplex. PA1 binding to a 524-bp DNA sequence with several cognate and noncognate (single bp mismatch) sites was also examined by quantitative footprinting, and SPR was performed on a noncognate hairpin treated with PA3 (see below).

6.4.2 Fluorescence Spectroscopy

Isotherms were observed using relatively large and easily reproducible decreases in TAMRA fluorescence intensity upon PA binding (10-50% depending on PA and dye location). Binding isotherms for PA2-4 are shown in Figure 6.3ABC (for the isotherm of PA1, see Ref. [22]). Unlabeled DNA was used in competition with dye-labeled DNA to evaluate potential dye-related artifacts (Figure 6. 3D; Ref. [22]); dyes were found not to interfere with interpretation (Table 6.1). Using this technique, we found that K_d values of $\leq$1 nM increased by ca. 100-fold when a single Py in PA1 was replaced by β to give PA2, and a similar increase in K_d occurred when a single internal Py was replaced in PA3 to give PA4.

6.4.3 Surface Plasmon Resonance

Using surface plasmon resonance, striking differences in behavior of PAs with and without an internal β residue are also easily seen in Figure 6.4AB, which features SPR sensorgrams for PA1 and PA2 binding to the cognate DNA sequence in Figure 6.2. The concentrations required to saturate the binding sites in PA1 are well below those required for PA2. Further, the kinetics, particularly for dissociation, are much slower for PA1 than PA2 (Table 6.1). PA3 binds even more strongly than PA1 to the same DNA hairpin; sensorgrams for PA3 binding are shown in Figure 6.4C. The sensorgrams for PA4 are very similar to those for PA2 and are not shown. The sensorgrams for PA1 and PA3 do not reach a steady-state plateau until almost saturation binding and were fit with a single-binding site kinetic model to determine both kinetics and equilibrium constants for binding [23-25]. As can be seen in the figures, the fits are excellent. The kinetics for PA2 and PA4 with DNA are too fast to fit with a kinetic model, but since they reach a steady-state, they were fit by using the steady-state RU values and

a single binding site model [23-25]. In this case, the fits are also excellent (not shown) and the results are in Table 6.1. These results clearly illustrate the detrimental effects of β-alanine substitution for DNA binding with these PAs, conditions and DNA sequences. The fact that no binding was detected by SPR between PA3 and a noncognate, single-bp mismatch DNA hairpin (Figure 6.4D) provides additional support for the highly selective DNA binding of PA1 and PA3, compounds lacking internal β residues.

At concentrations of 10 nM and higher, all four PAs bind rapidly to DNA in SPR experiments and reach a steady state plateau. The dissociation rate constants, however, are much smaller for PA1 and PA3 than for PA2 and PA4. This observation suggests that all of the polyamides rapidly locate their DNA binding sites and associate with the minor groove. Once bound, however, PA1 and PA3 form exceptionally tight complexes that dissociate very slowly from DNA. In these cases, in contrast to the systems mentioned above, the flexible β linker prevents tight binding of PAs in comparison to equivalently substituted pyrrole derivatives. Furthermore, the more rigid species PA1 and PA3 show excellent discrimination against single-base pair mismatches (Figure 6.3CD), again in contrast to prior reports of similar molecules [32]. Table 6.1 illustrates the excellent agreement found between fluorescence and SPR measurements for both the tight binding PA1 and PA3 and the weaker-binding PA2 and PA4, providing validation for both experimental approaches.

6.4.4 Quantitative DNase I Footprinting

To compare single site binding affinities of PA1 and PA2 for a linear dsDNA molecule larger than the above DNA hairpins, a 120 bp duplex with a single PA1/2 binding site was generated. The central part of the sequence is identical to the stem portion of the above hairpin

used for fluorescence and SPR studies (see Section 2.5). The flanking sequences were derived from a portion of the HPV16 genome that we predicted to contain no additional binding sites for PA1 or PA2. Quantitative DNase I footprinting on this duplex yields K_ds of 1.1 ± 0.2 nM [22] and 80 ± 8 nM for PA1 and PA2, respectively. These data agree remarkably well with the results obtained by other methods, providing additional evidence that K_ds obtained with the relatively small DNA hairpins are indicative of affinities observed in larger DNA contexts.

To determine how PA1 binds its target in an even larger sequence context, we studied PA1 binding to an AT-rich 524 bp duplex from the HPV16 genome using DNase I footprinting and capillary electrophoresis [27, 33]. As summarized in Table 6.2 and illustrated in Figure 6.5, PA1 binds to three cognate sites with high affinity. This provides further confirmation of the aforementioned observations involving a small DNA hairpin, and indicates that PA1 binds its target avidly and specifically within the larger context of a long DNA sequence. Binding at single bp mismatch sites in the 524-mer was so weak as to be largely unobserved at [PA] ≤ 60 nM, with resulting noncognate K_d values conservatively estimated at ≥ 60-100 nM.

While anomalies in DNA binding of β-substituted polyamides have been mentioned [20, 21], no quantitative details of such behavior have been previously reported to our knowledge. Instead of finding improved binding strength and specificity upon incorporating β for Py in an 8-ring polyamide [32], we found the opposite. The binding results reported here took on additional significance for us because of our use of PA molecules as anti-human papillomavirus (HPV) agents in collaboration with NanoVir, LLC. Screening libraries of compounds with internal β/Py substitutions shuffled throughout the sequence targets the same cognate binding site, but produces both active and inactive anti-HPV compounds [18, 34]. The present report helps guide future PA design and structure-activity studies for our own and other labs. In particular, the

effect of incorporating β-alanine into PA designs is not straightforward, and the standard binding rules can be insufficient to predict DNA binding strength or selectivity. The same may hold true for predicting polyamide activity in cells, tissue culture and whole animals, and this hypothesis is under current investigation.

6.5 Acknowledgments

We thank NIH (NIAID AI083803 to JKB, NIAID AI064200 to WDW) and NanoVir, LLC for financial support and the Danforth Plant Science Center for HRMS (obtained with an LTQ Orbitrap Velos purchased under grant NSF-DBI 0922879); JKB thanks J. J. Shieh for SPR work on a prior COX-2 project. JKB is part owner of NanoVir, LLC

6.6 References

[1] Muzikar K.A., Nickols N.G., Dervan P.B., Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression, Proc. Natl. Acad. Sci. USA 106 (2009) 16598-16603.

[2] Harki D.A., Satyamurthy N., Stout D.B., Phelps M.E., Dervan P.B., In vivo imaging of pyrrole-imidazole polyamides with positron emission tomography, Proc. Natl. Acad. Sci. USA 105 (2008) 13039-13044.

[3] Shinohara K.-I., Bando T., Sugiyama H., Anticancer activities of alkylating pyrrole-imidazole polyamides with specific sequence recognition, Anti-Cancer Drugs 21 (2011) 228-242.

[4] Ueno T., Fukuda N., Tsunemi A., Yao E.-H., Matsuda H., Tahira K., Matsumoto T., Matsumoto K., Matsumoto Y., Nagase H., Sugiyama H., Sawamura T., A novel gene silencer, pyrrole-imidazole polyamide targeting human lectin-like oxidized low-density lipoprotein receptor-1 gene improves endothelial cell function, J. Hyperten. 27 (2009) 508-516.

[5] Sasaki S., Bando T., Minoshima M., Shinohara K.-i., Sugiyama H., Sequence-specific alkylation by Y-shaped and tandem hairpin pyrrole-imidazole polyamides, Chem. Eur. J. 14 (2008) 864-870.

[6] Dervan P.B., Edelson B.S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides, Curr. Opin. Struct. Biol. 13 (2003) 284-299.

[7] Dervan P.B., Doss R.M., Marques M.A., Programmable DNA binding oligomers for control of transcription, Current medicinal chemistry: Anti-cancer agents 5 (2005) 373-387.

[8] Dose C., Farkas M.E., Chenoweth D.M., Dervan P.B., Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit, J. Amer. Chem. Soc. 130 (2008) 6859-6866.

[9] Warren C.L., Kratochvil N.C., Hauschild K.E., Foister S., Brezinski M.L., Dervan P.B., Phillips G.N., Jr., Ansari A.Z., Defining the sequence-recognition profile of DNA-binding molecules, Proc. Natl. Acad. Sci. USA 103 (2006) 867-872.

[10] Farkas M.E., Tsai S.M., Dervan P.B., Alpha-diaminobutyric acid-linked hairpin polyamides, Bioorg. Med. Chem. 15 (2007) 6927-6936.

[11] Nishijima S., Shinohara K., Bando T., Minoshima M., Kashiwazaki G., Sugiyama H., Cell permeability of Py-Im-polyamide-fluorescein conjugates: Influence of molecular size and Py/Im content, Bioorg. Med. Chem. 18 (2010) 978-983.

[12] Minoshima M., Bando T., Shinohara K., Kashiwazaki G., Nishijima S., Sugiyama H., Comparative analysis of DNA alkylation by conjugates between pyrrole-imidazole hairpin polyamides and chlorambucil or seco-CBI, Bioorg. Med. Chem. 18 (2010) 1236-1243.

[13] Maeshima K., Janssen S., Laemmli U.K., Specific targeting of insect and vertebrate telomeres with pyrrole and imidazole polyamides, EMBO J. 20 (2001) 3218-3228.

[14] Janssen S., Durussel T., Laemmli U.K., Chromatin opening of DNA satellites by targeted sequence-specific drugs, Mol. Cell. 6 (2000) 999-1011.

[15] Babu B., Liu Y., Plaunt A., Riddering C., Ogilvie R., Westrate L., Davis R., Ferguson A., Mackay H., Rice T., Chavda S., Wilson D., Lin S., Kiakos K., Hartley J.A., Lee M.,

Design, synthesis and DNA binding properties of orthogonally positioned diamino containing polyamide f-IPI, Biochem. Biophys. Res. Comm. 404 (2011) 848-852.
[16] Sielaff A., Cooper A., Mackay H., Brown T., O'Hare C., Kluza J., Kotecha M., Le M., Hochhauser D., Hartley J.A., Lee M., Binding of f-PIP and JH-37 to the inverted CCAAT box-2 of the topoisomerase IIa promoter, Abstracts of Papers, 233rd ACS National Meeting, Chicago, IL, United States, March 25-29, 2007, p. 134.
[17] Reddy P.M., Dexter R., Bruice T.C., DNA sequence recognition in the minor groove by hairpin pyrrole polyamide-Hoechst 33258 analogue conjugate, Bioorg. Med. Chem. Lett. 14 (2004) 3803-3807.
[18] Edwards T.G., Koeller K.J., Slomczynska U., Fok K., Helmus M., Bashkin J.K., Fisher C., HPV episome levels are potently decreased by pyrrole-imidazole polyamides, Antiviral Res. 91 (2011) 177-186.
[19] Ehley J.A., Melander C., Herman D., Baird E.E., Ferguson H.A., Goodrich J.A., Dervan P.B., Gottesfeld J.M., Promoter scanning for transcription inhibition with DNA-binding polyamides, Mol. Cell. Biol. 22 (2002) 1723-1733.
[20] Dervan P.B., Urbach A.R., The importance of β-alanine for recognition of the minor groove of DNA, Essays Contemp. Chem. (2001) 327-339.
[21] Turner J.M., Swalley S.E., Baird E.E., Dervan P.B., Aliphatic/Aromatic Amino Acid Pairings for Polyamide Recognition in the Minor Groove of DNA, J. Am. Chem. Soc. 120 (1998) 6219-6226.
[22] Dupureur C.M., Bashkin J.K., Aston K., Koeller K.J., Gaston K.R., He G., Fluorescence assay of polyamide-DNA interactions, Anal. Biochem. 423 (2012) 178-183.
[23] Liu Y., Wilson W.D., Quantitative Analysis of Small Molecule–Nucleic Acid Interactions with a Biosensor Surface and Surface Plasmon Resonance Detection, in: Fox K.R. (Ed.), Methods Mol. Biol., Springer, NY, 2010, pp. 1-23.
[24] Myszka D.G., Improving biosensor analysis, J. Mol. Recog. 12 (1999) 279-284.
[25] Nguyen B., Tanious F.A., Wilson W.D., Biosensor-surface plasmon resonance: Quantitative analysis of small molecule-nucleic acid interactions, Methods 42 (2007) 150-161.
[26] Trauger J.W., Dervan P.B., Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes, Methods Enzymol. 340 (2001) 450-466.
[27] Mitra S., Shcherbakova I.V., Altman R.B., Brenowitz M., Laederach A., High-throughput single-nucleotide structural mapping by capillary automated footprinting analysis Nucl. Acids Res. 36 (2008) e63.
[28] Baird E.E., Dervan P.B., Solid Phase Synthesis of Polyamides Containing Imidazole and Pyrrole Amino Acids, J. Am. Chem. Soc. 118 (1996) 6141-6146.
[29] Phillion D.P., Crowley K.S., Bashkin J.K., Schweitzer B.A., Burnett B.L., Woodard S.S., Polyamide modulators of COX2 transcription, Pharmacia & UpJohn Company, published PCT International Patent Appln., WO 03/040337, 2003, 40 pp.
[30] Dickinson L.A., Trauger J.W., Baird E.E., Dervan P.B., Graves B.J., Gottesfeld J.M., Inhibition of Ets-1 DNA binding and ternary complex formation between Ets-1, NF-kappaB, and DNA by a designed DNA-binding ligand, J. Biol. Chem. 274 (1999) 12765-12773.
[31] Garvie C.W., Pufall M.A., Graves B.J., Wolberger C., Structural analysis of the autoinhibition of Ets-1 and its role in protein partnerships, J. Biol. Chem. 277 (2002) 45529-45536.

[32] Wang C.C., Ellervik U., Dervan P.B., Expanding the recognition of the minor groove of DNA by incorporation of beta-alanine in hairpin polyamides, Bioorg. Med. Chem. 9 (2001) 653-657.
[33] Wilson D.O., Johnson P., McCord B.R., Nonradiochemical DNase I footprinting by capillary electrophoresis, Electrophoresis 22 (2001) 1979-1986.
[34] Fisher C., Bashkin J.K., Crowley K.S., Sverdrup F.M., Garner-Hamrick P.A., Phillion D.P., Polyamide Compositions and Therapeutic Methods for Treatment of Human Papilloma Virus, Pharmacia & UpJohn Company, published PCT International Patent Appln., WO 05/03382, 2005, 53

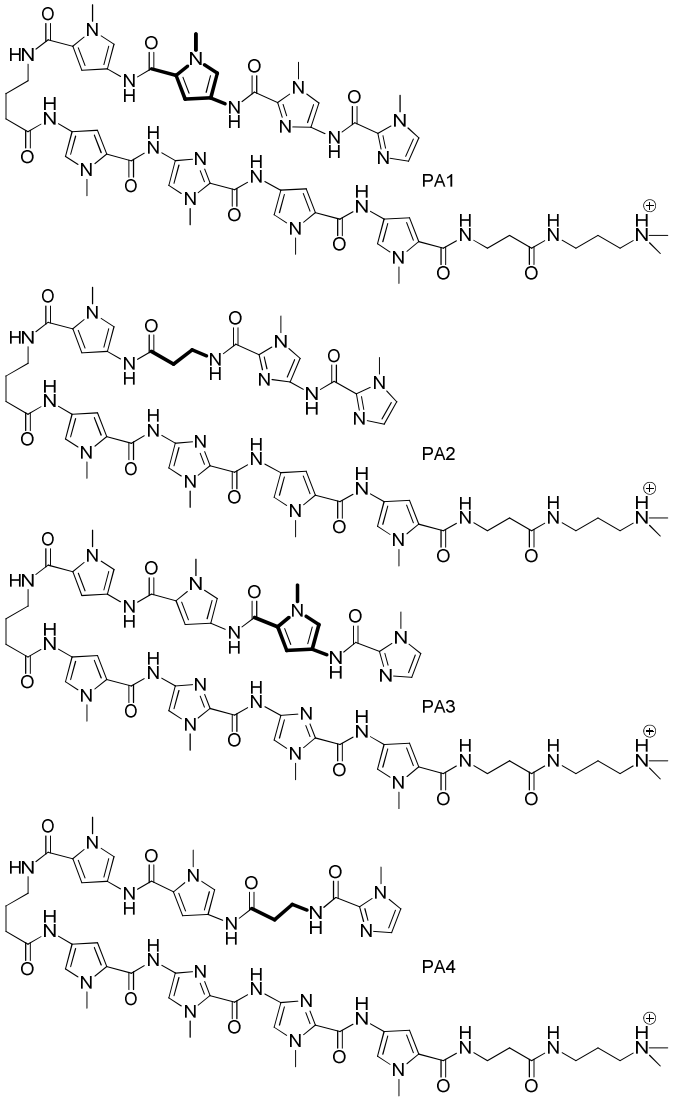


Figure 6.1 Structures of compounds PA1-PA4. They are shown as monocations for clarity; they are isolated as tetracations. Bold bonds indicate the Py and β pairs compared in this study.

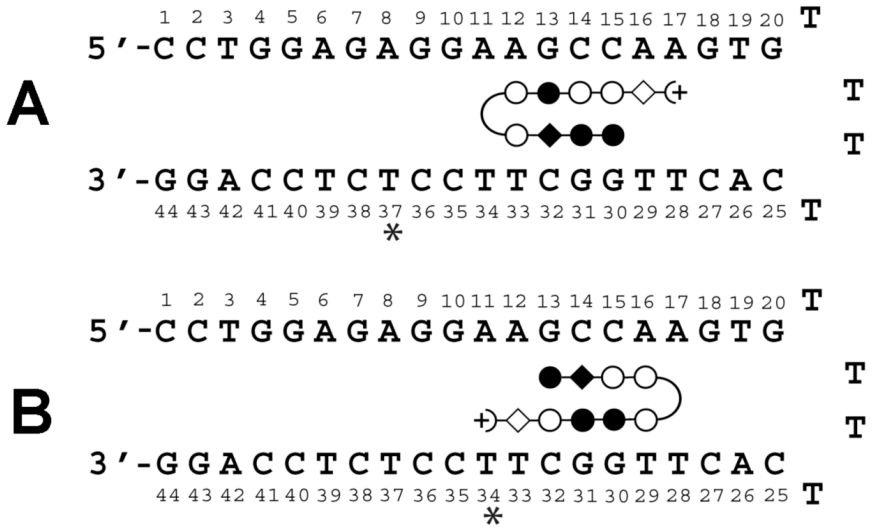


Figure 6.2 Map of predicted polyamide-DNA interactions: (**A**) PA1/2 and (**B**) PA3/PA4 bound to hairpin DNA. Open circle: Py; filled circle: Im; open diamond: β; filled diamond: Py (PA1, PA3) or β (PA2, PA4); (+: Dp; curved line: γ; *: position of TAMRA dye when used. A 5'-Biotin conjugate of the DNA without dye labels was used for SPR.

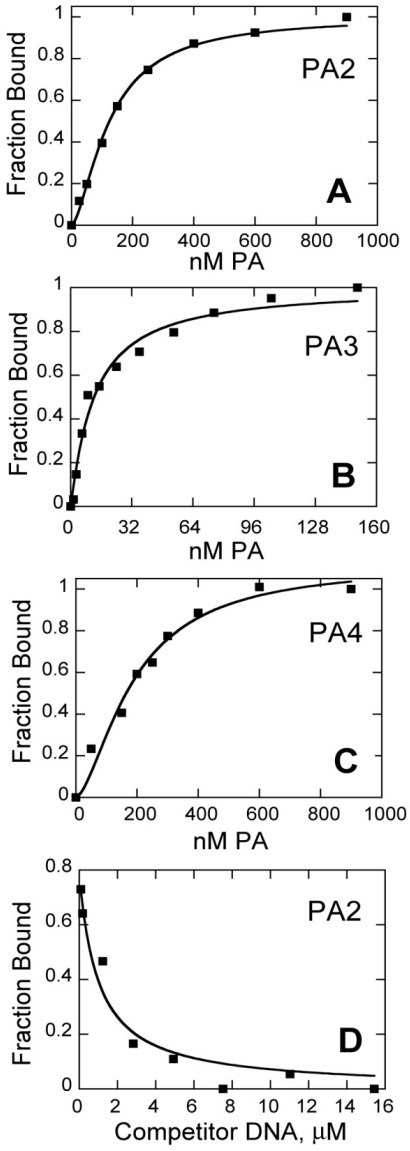


Figure 6.3 Polyamide-DNA binding as observed via fluorescence spectroscopy. Binding isotherms for PA2 (**A**), PA3 (**B**), and PA4 (**C**) toward TAMRA-labeled DNA hairpin as described in the text. (**D**) Sample competition experiment between an unlabeled DNA hairpin and a labeled DNA hairpin bound to PA2. Conditions: 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, pH 7.4, 25 °C.

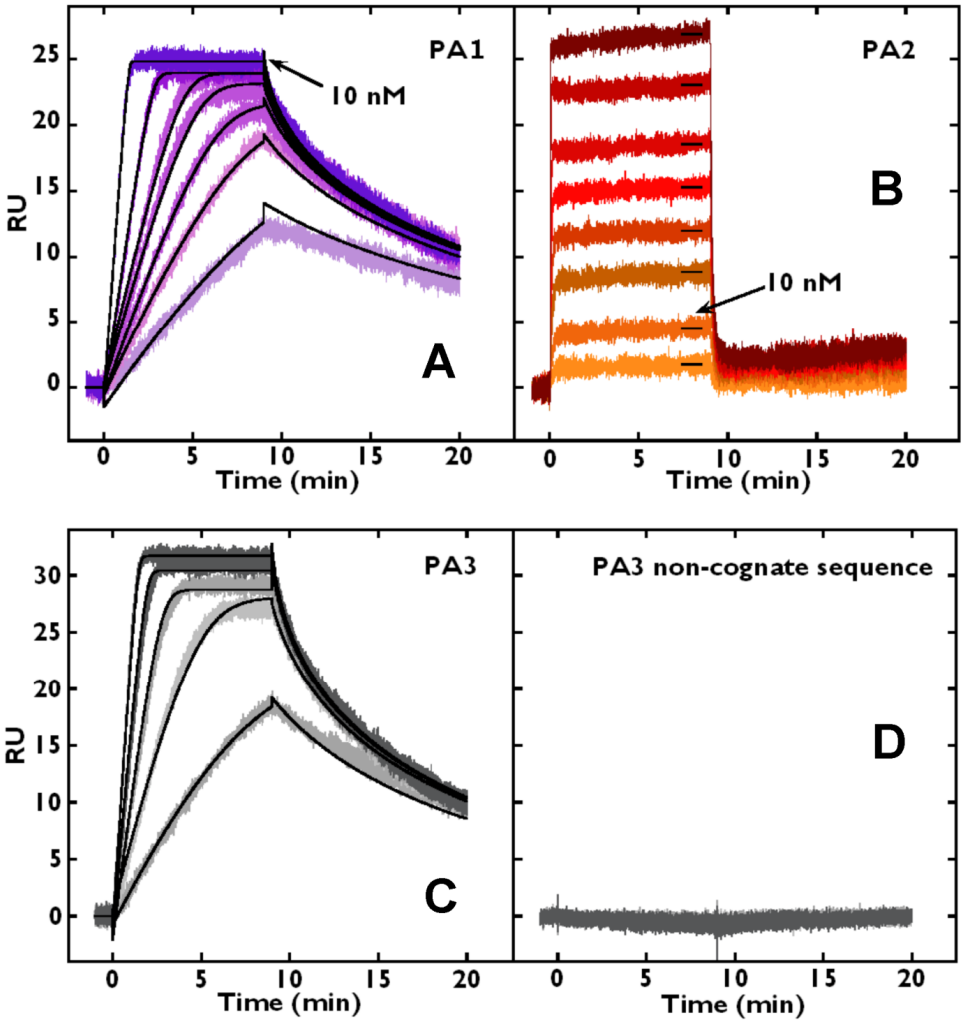


Figure 6.4 Polyamide-DNA Binding as observed via SPR. Sensorgrams for the interactions of PA1 (**A**) and PA2 (**B**) with the hairpin DNA shown in Figure 6.2 labeled with 5'-biotin. Individual sensorgrams represent responses at different PA concentrations: concentrations for PA1 are 1.0, 1.5, 2.0, 2.5, 3.0, 5.0 and 10.0 nM; concentrations for PA2 are 1.0, 10, 20, 40, 60, 100, 200 and 400 nM. (**C**) Sensorgram for the interactions of PA3 with the cognate hairpin DNA from Figure 6.2 (**D**) the noncognate hairpin DNA, sequence 5'-Biotin-CCTTGGAGAGTTTTCTCTCCAAGG-3', with the hairpin loop underlined. The individual sensorgrams represent responses at different PA concentrations. The concentrations for PA3 are 1.0, 2.0, 4.0, 6.0 and 8.0 nM in both panels. Kinetic and steady-state fits used BIAcore T100 Evaluation Software.

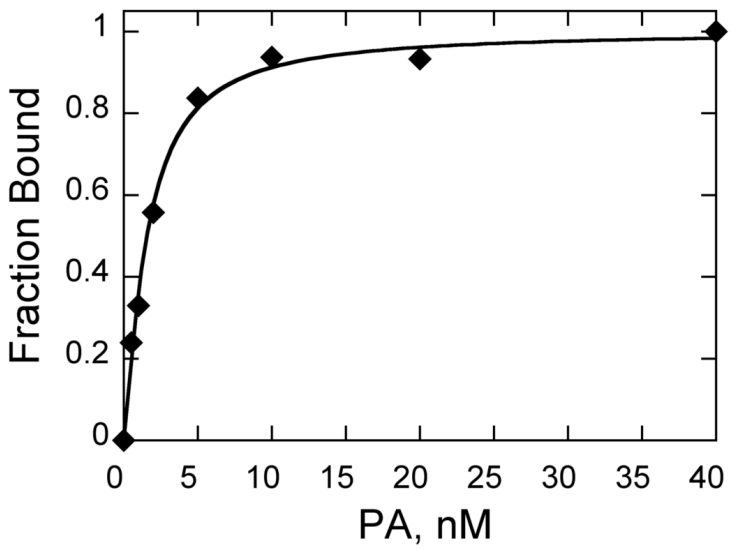


Figure 6.5 PA1-DNA Binding via Quantitative Footprinting and CE. Sample isotherm of PA1 binding an AT rich 524 DNA duplex as detected by DNase I footprinting and capillary electrophoresis; conditions: 0.2 nM duplex; otherwise as per text. Data fit to a K_d of 1.6 nM.

Table 6.1 Data for DNA binding of PA1-PA4 to DNA hairpin by fluorescence (FL) and SPR.

Cmpd.	Sequence	K_d/FL (nM)[a]	K_d/FL competition (nM)[b]	K_d/SPR (nM)	k_a/SPR ($M^{-1}s^{-1}$)	k_d/SPR (s^{-1})
PA1	dImIm**Py**PyγPyImPyPyβDp	2.3 ± 0.3[d]	1.1 ± 0.2[d]	1.12 ± 0.21	5.3×10^7	0.059
PA2	dImIm**β**PyγPyImPyPyβDp	170 ± 40	325 ± 150	83.4 ± 8.9	tf[c]	tf
PA3	dIm**Py**PyPyγPyImImPyβDp	11.0 ± 1.0[d]	≤ 1.0	0.71 ± 0.14	1.2×10^7	0.0085
PA4	dIm**β**PyPyγPyImImPyβDp	120 ± 15	70 ± 30	106 ± 11	tf	tf

[a]C-5 TAMRA dye substitution at T37 for PA1-2 and T34 for PA3-4. [b]From competition of unlabeled DNA with dye-labeled DNA. [c]tf: too fast to determine, most of the interaction occurs during filling of the flow cells and the dissociation occurs during filling with buffer. [d]Reproduced from Ref. [22]

Table 6.2. Sites and K_d values for PA1 bound to a 524 bp DNA[a].

Polyamide binding site	**K_d (nM)**
5'-AGGCAT-3'	1.6±0.2
5'-AAGCCA-3'	3.3±1
5'-ATGCCA-3'	2.6±1
5'-AGG**T**AT-3'	60-100
5'-AGG**T**AA-3'	>100
5'-TGGC**C**T-3' (two sites)	>60

[a]Determined by quantitative DNase I footprinting; mismatches in bold.